High Points of a Gaussian Free Field and a Gaussian Membrane Model and Limit Shape of Young Diagrams for Random Permutations

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Alessandra Cipriani

aus

Italien

Promotionskomitee

Prof. Dr. Erwin Bolthausen, Vorsitz
Prof. Dr. Ashkan Nikeghbali
Prof. Dr. Daniel Ueltschi

Zürich, 2014

Abstract

This thesis consists of two distinct parts. In the first part, we treat two different models of Gaussian Fields: one is the membrane model at its critical dimension, of which we establish the Hausdorff dimension of its high points, and the other one is the continuum Gaussian Free Field in dimension 4, of which we determine the Hausdorff dimension of the thick points and prove they constitute the support of the 4-dimensional Liouville Quantum Gravity measure. In the second part, we deal with random permutations and show the limit shape for Young diagrams under a so-called conservative measure on the set of permutations on n objects.

Zusammenfassung

Diese Dissertation besteht aus zwei verschiedenen Hauptteilen. Im ersten Teil betrachten wir zwei verschiedene Modelle von Gausschen Felden: Eines ist das Membranmodell in seiner kritischen Dimension, von dem wir die Hausdorff-Dimension der hohen Punkte bestimmen, und das andere ist das 4-dimensionale "Continuum Gaussian Free Field", von dem wir die Hausdorff-Dimension der hohen Punkte bestimmen und zeigen, dass sie den Träger des Liouville-Quantum-Gravitationsmasses bezeichnen. Im zweiten Teil behandeln wir Zufallspermutationen und zeigen das Grenzprofil von Young Diagrammen bezüglich eines sogenannten konservativen Masses auf der Menge der Permutationen von n Objekten.

What thou lovest well remains, the rest is dross
What thou lov'st well shall not be reft from thee
What thou lov'st well is thy true heritage
Ezra Pound, Pisan Cantos, LXXXI.

Acknowledgments

I regard myself as a very lucky person for having so many people to thank. My PhD journey has been intense and hard at times, but looking back now it was worth to the last bit – and possibly these long acknowledgments are part of the reward.

Many thanks to my advisor Erwin Bolthausen for having given me so many opportunities during my doctoral studies, and for having given me the freedom to choose the topics I enjoyed the most. I have never felt under pressure or judged in any way, and I am thankful for this.

I likewise thank Ashkan Nikeghbali and Daniel Ueltschi, who kindly agreed to referee my thesis and go carefully through it.

I want to thank Dirk Zeindler for having taken care of correcting all my mistakes during our work, and for his constant motivation towards math which has been a drive to become a better researcher. I have benefitted a lot from our collaboration and I am glad he chose me as collaborator.

I had the great honor of working with Rajat Hazra: not only is he an outstanding mathematician, but also his modesty and wisdom have been a great reference point to me and I feel deeply indebted for his heartfelt support in any occasion; I hope one day to be able to advise someone the same way he did with me at the ETH smoking zone.

Noemi Kurt and Luca Avena deserve also a special mention. They were always there whenever I needed help for the toughest math questions and if I had doubts on whether I could be at their stage one day. Their sensible suggestions meant and still mean a lot to me, not only in the academic world.

When I first arrived in Zurich I was told that "the working environment there is awesome". I couldn't believe it at first, but it was true. I have met great friends at the institute who have become my Swiss family. I would like to thank in particular Elisabetta Chiodaroli for having been a compass when I was about to get lost; Sara Daneri for being a true coach; Kim Dang, who always supported me with kind words; Anne Feidt, for having written together with me a book through emails, for all our laughters, research sessions, lunches together, mega-whines and literary reviews; Felix Fontein, who helped me not to get mad when things went crazy; Marta Pittavino, for looking up at me as a role model (we have a book to write together too) and for her wise advice; Philipp Thomann, for simply being a true older brother; Alberto Vezzani, for everytime he asks the most simple question: "How are you?" he means it.

My office feels like home now, and my officemates managed to make it so since the very first days. A big thank-you to Cyril Marzouk, Thomas Preu, Emanuele Spadaro, Julia Storm and Tamara Widmer. Erich Baur was more of a twin since we started together, and I am very glad for having found him to come alongside during my PhD (everything is doable, Erich!). I hope we will manage to use our help line for many years to come.

I also have to mention many more people with whom I had a great time: special thanks to Stefania Colangelo, Leif Döring, Iván Contreras, Andrea Ferraguti, Nicoletta Gabrielli, Steven Geinitz, Elisa Gorla, Mahir Hadzić, Jonas Hirsch, Felice Manganiello, Anna Mätzener, Giacomo Micheli, Tommaso Mingazzini, Felix Rubin, Marko Serić, Laura Spinolo, Luca Spolaor, Eva Tudor, Marko Živković; the "Berliners" who were so friendly during my stay there, in particular Marcel Ortgiese and Marcia Ramalho.

My thanks to Dima Ioffe for having invited me to the Technion and to Nick Crawford for working so intensely with me during my stay there as I had just started my PhD. I learnt a great deal in only one month in Israel and I still carry it with me today.

I have a big debt towards Francesca Collet, Anna Levitt and Wioletta Ruszel for their wonderful words of encouragement when I most needed them, and which I keep on playing in my head when I am unsure on what to do.

Herzlichen Dank an Daniela Secchi, die mich in Deutsch unterrichtete. 4 Jahre später bin ich von den Ereignissen beeindruckt.

Vielen vielen Dank auch an Franziska Robmann, die sie sich immer Mühe gegeben hat, um meine bürokratischen Probleme zu lösen und mein Leben zu vereinfachen.

Thanks to Elisabetta Candellero, Hubert Lacoin and Michele Salvi, with whom any conference becomes such an enriching and fun experience.

Of my two lifes, as I call them, I can't forget about the second one too. All of my friends who stayed in Italy when I moved actually never left me, and I can't remember a single time when they were not ready to share something with me. Edina, Giulia, Laura, Mariachiara, Nicolò: we are grown up now and I am pleased to say we have done it together. Thanks to awesome Mrs. President Laura for her great friendship and to Diego for all the good compliments he paid to me and that I not sure I deserve; with Paola I was trying to pen down how much I owe her, but I guess this will never be possible. Looking back to 2013 will probably be the only gift I can eventually give her.

And finally to my parents Grazia and Roberto because they love me even when we do not agree, and to my brother for being the first person I turn to during my good and bad days.

Thank you so much.

Contents

Part I.

Random interfaces and Gaussian Fields

Chapter 1.

The Membrane Model

1.1. Random interfaces

The field of random interfaces has been widely studied in statistical mechanics. These interfaces are described by a family of random variables indexed by the d-dimensional integer lattice, which are considered as a height configuration, namely they indicate the height of the interface above a reference hyperplane. More formally, we call any collection of real numbers

$$\left\{\varphi_x : x \in \mathbb{Z}^d\right\},$$

$d \geq 1$, an interface. We identify the family $\{\varphi_x\}$ with the (graph of the) mapping

$$\varphi : \mathbb{Z}^d \to \mathbb{R}$$

s. t. $\varphi(x) = \varphi_x$. A probability measure on the set of interface configurations can be

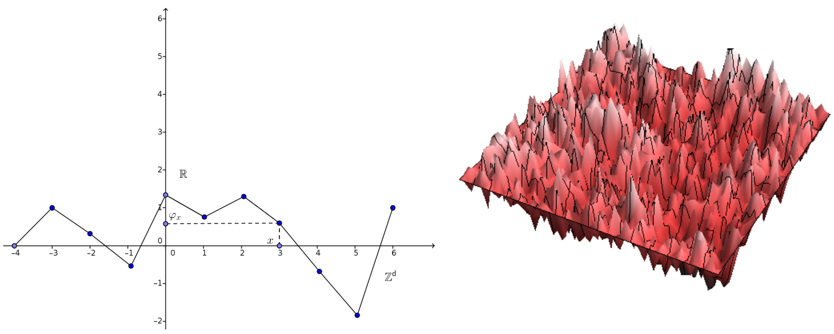

Figure 1.1.: An example of interface in $\mathbb{Z}$ (Gaussian random walk) and $\mathbb{Z}^2$ (Discrete Gaussian Free Field).

introduced as follows. Let $\Omega := \mathbb{R}^{\mathbb{Z}^d}$ be endowed with the product topology. Let Λ be a finite subset of $\mathbb{Z}^d$. We fix a configuration ψ to be the boundary condition.The probability of a configuration depends on its energy (the Hamiltonian $H^{\psi}_{\Lambda}(\varphi)$. The corresponding distribution of φ is then the following (formal) quantity:

$$P^{\psi,\beta}_{\Lambda}(\mathrm{d}\varphi) = \frac{\exp\left(-\beta H^{\psi}_{\Lambda}(\varphi)\right)}{Z^{\psi,\beta}_{\Lambda}} \prod_{x\in\Lambda} \mathrm{d}\varphi_x \prod_{x\in\Lambda^c} \delta_{\psi_x}(\mathrm{d}\varphi_x).$$

$\beta \geq 0$ is called the inverse temperature, $\mathrm{d}\varphi_x$ is the one-dimensional Lebesgue measure, δ_{ψ_x} is the Dirac mass at ψ_x and $Z^{\psi,\beta}_{\Lambda}$ is a normalizing constant. Let us see some examples of Hamiltonians which we will consider in this thesis.

Definition 1.1. *A* gradient model *(or ∇–model) is a random interface model in the context we just described, where the Hamiltonian is given by*

$$H^{\psi}_{\Lambda}(\varphi) = \frac{1}{2}\sum_{x,y\in\Lambda} p_{x,y} V(\varphi_x - \varphi_y) + \sum_{x\in\Lambda, y\notin\Lambda} p_{x,y} V(\varphi_x - \varphi_y). \tag{1.1}$$

$V : \mathbb{R} \to \mathbb{R}$ is assumed to be an even convex function with $V(0) = 0$ and $p_{x,y}$ is the transition matrix of a random walk in the lattice.

If the random walk is for example of finite range (even though one can find more general conditions, see [79, 41]) then $P^{\psi,\beta}_{\Lambda}$ is a probability measure. The most studied model falling into such category is the Discrete Gaussian Free Field (DGFF), also called harmonic lattice or harmonic crystal, whose Hamiltonian is given by (1.1) when $V(x) = x^2$. Then we have the alternative representation

$$H^{\psi}_{\Lambda}(\varphi) = \frac{1}{2}\sum_{x,y\in\Lambda} \varphi_x (\mathbb{I} - P)_{\Lambda}(x,y)\varphi_y + \sum_{x\notin\Lambda} m^{\psi}_x \varphi_x$$

for some coefficients $m^{\psi}_x \in \mathbb{R}$. Here $(\mathbb{I} - P)_{\Lambda}(x,y) = \delta_{x,y} - p_{x,y}$. Hence the DGFF represents a multivariate centered Gaussian random variable whose covariance matrix satisfies the relation

$$\Gamma_{\Lambda}(x,y) := \mathrm{cov}(\varphi_x,\, \varphi_y) = \mathbb{E}^x\left(\sum_{n=1}^{\tau_{\partial\Lambda}-1} \mathbb{1}_{\{S_n = y\}}\right), \tag{1.2}$$

where $\mathbb{E}^x$ is the law of a standard random walk (SRW) $(S_n)_{n\in\mathbb{N}}$ started at $x \in \mathbb{Z}^d$ and $\tau_{\partial V_N}$ is the first exit time from Λ:

$$\partial\Lambda := \{x \in \Lambda^c : \mathrm{dist}(x,\, V_N) = 1\},$$

$$\tau_{\Lambda} := \min\{n \geq 0 : S_n \notin \Lambda\}.$$

This is what is usually called the *random walk representation* of the covariances ([59]). Note that here

$$p_{x,y} = \frac{1}{2d}\,\mathbb{1}_{\{x\sim y\}}$$

and more specifically the Hamiltonian has the equivalent form

$$H(\varphi) = \frac{1}{2d}\sum_x |\nabla \varphi_x|^2,$$

with the vector $\nabla\varphi_x = (\varphi_x - \varphi_{x+\vec{e}_1}, \ldots, \varphi_x - \varphi_{x+\vec{e}_d})$. With the gradient representation it is easier to see that the DGFF favors configurations whose height is approximately constant, since each point is roughly at height averaged among its nearest neighbors. This statement has in fact a more precise formulation using discrete differential operators:

Definition 1.2. Δ *is the* discrete Laplacian *which we define as the matrix* $\mathbb{I} - P =: -\Delta$, *in particular*

$$\Delta(x,y) = \begin{cases} -1 & x = y \\ \frac{1}{2d} & x \sim y \\ 0 & \text{otherwise} \end{cases}.$$

Alternatively one can define it as a differential operator acting on functions $f : \mathbb{Z}^d \to \mathbb{R}$ *at a point* $x \in \mathbb{Z}^d$

$$\Delta f(x) = \frac{1}{2d}\sum_{i=1}^{d} f(x+\vec{e}_i) + f(x-\vec{e}_i) - f(x).$$

One can convince himself quickly that Δ is symmetric and positive definite. The link to partial differential equations becomes then evident when one notices that for $x \in \Lambda$ the function $\Gamma_\Lambda(x,\cdot)$ represents the unique solution of the discrete boundary value problem

$$\begin{cases} \Delta\Gamma_\Lambda(x,y) = \delta_{xy} & y \in \Lambda \\ \Gamma_\Lambda(x,y) = 0 & y \in \partial\Lambda. \end{cases} \tag{1.3}$$

Practically Γ_Λ is the Green's function of the discrete Laplacian with Dirichlet boundary conditions on Λ. This yields a number of interesting properties, namely for all x, $y \in \mathbb{Z}^d$

- symmetry: $\Gamma_\Lambda(x,y) = \Gamma_\Lambda(y,x)$,
- monotonicity: if $\Lambda' \subseteq \Lambda \subseteq \mathbb{Z}^d$, $\Gamma_{\Lambda'}(x,y) \le \Gamma_\Lambda(x,y)$.

It is also possible to prove the existence of the infinite volume limit which stems from the transience of the SRW in dimension at least 3. Here we consider the limit of finite-volume distributions taken with respect to the weak topology on the space of probability measures.

Lemma 1.3. *The infinite volume measure of the harmonic lattice exists for $d \geq 3$. It is the centered Gaussian field with covariance matrix $\Gamma(x, y) = \mathbb{E}^x \left(\sum_{n=1}^{+\infty} \mathbb{1}_{\{S_n = y\}} \right)$.*

The existence of the limit allows us to define a probability on $\mathbb{Z}^4$, a fact which is in general not doable a priori (see for example [38]). The proofs of these facts can be found for example in the comprehensive reference [59] and rely on the tight connection to the SRW.

For the rest of this section $V_N := [-N, N]^d \cap \mathbb{Z}^d$, $N > 0$ fixed, will play the role of the finite volume on which we consider our probability distribution. We will denote as $\Delta_N = (\Delta_N)(x,y) := \Delta(x,y)_{x,y \in V_N}$.

1.2. The membrane model

The Membrane Model is a Gaussian multivariate random variable whose Hamiltonian depends on the mean curvature of the interface, in particular favors configurations whose curvature is approximately constant. The study of such interface was firstly undertaken by Sakagawa in [71]; we are aware of the contributions of Kurt ([57], [58]) regarding also a phenomenon called *entropic repulsion* in dimension 4, and on the so-called *pinning* and *wetting* phase transitions which were considered in dimension 1 by [17, 18]. A mixed ∇- and membrane model is the object of study of [16].

Definition 1.4. *The Membrane model with 0 boundary conditions outside V_N is a Gaussian field which can be equivalently seen as:*

(a) a the random interface model whose distribution is given by

$$\mathbf{P}_N(\mathrm{d}\varphi) = \frac{1}{Z_N} \exp\left(-\frac{1}{2} \sum_{x \in \mathbb{Z}^d} (\Delta \varphi_x)^2 \right) \prod_{x \in V_N} \mathrm{d}\varphi_x \prod_{x \in \partial_2 V_N} \delta_0(\mathrm{d}\varphi_x), \tag{1.4}$$

where Δ is the discrete Laplacian, $\partial_2 V_N := \{y \in V_N^c : \operatorname{dist}(y, V_N) \leq 2\}$ and Z_N is the normalizing constant.

(b) By re-summation, the law $\mathbf{P}_N$ of the field is the law of the centered Gaussian field on V_N with covariance matrix

$$G_N(x,y) := \mathbf{Cov}_N(\varphi_x, \varphi_y) = (\Delta_N^2)^{-1}(x,y).$$

Here, $\Delta_N^2(x,y) = \Delta^2(x,y) 1_{\{x,y \in V_N\}}$ is the Bilaplacian or biharmonic operator with 0-boundary conditions outside V_N. It can be seen as the (symmetric and positive definite) *matrix given by*

$$\Delta_N^2(x,y) = \begin{cases} 1+\frac{1}{2d} & x=y,\, x\in V_N \\ -\frac{1}{d} & x\sim y,\, x,\, y\in V_N \\ \frac{1}{4d^2} & |x-y|=2,\, x,\, y\in V_N \\ \frac{1}{2d^2} & |x-y|=\sqrt{2},\, x,\, y\in V_N \\ 0 & \text{otherwise} \end{cases}.$$

(c) *The model is a centered Gaussian field on V_N whose covariance matrix G_N satisfies, for $x \in V_N$,*

$$\begin{cases} \Delta^2 G_N(x,y) = \delta_{xy}, & y\in V_N \\ G_N(x,y)=0, & y\in \partial_2 V_N. \end{cases}$$

Remark 1.5. *It is crucial to notice that*

$$\Delta_N^2 \neq (\Delta_N)^2. \tag{1.5}$$

1.2.1. Phase transition of the Membrane model

As for many other models in statistical physics, the Membrane model features a *phase transition*, i. e. a sharp difference in its behavior when the dimension changes.

Lemma 1.6 ([56]). *The Membrane model is subcritical for $d \le 3$, critical at $d = 4$ and supercritical at $d \ge 5$.*

This transition entails:

(a)

$$\mathbf{Var}_N(\varphi_0) = \begin{cases} c_d N^{4-d} + o\left(N^{4-d}\right) & d\le 3 \\ c_4 \log N + O(1) & d=4 \\ c_d + O\left(N^{4-d}\right) & d\ge 5 \end{cases}$$

(we will state this result more precisely for $d \ge 4$ in Subsection 1.4.1);

(b) for $d \ge 5$ the infinite volume Gibbs measure P exists [56, Prop. 1.2.3] and is the law of the centered Gaussian field with covariance matrix

$$G(x,y) = \Delta^{-2}(x,y) = \mathbb{E}^{x,y}\left(\sum_{n,m=1}^{+\infty} \mathbb{1}_{\{S_n=T_m\}}\right)$$

with $(S_n)_{n\in\mathbb{N}}$, $(T_m)_{m\in\mathbb{N}}$ two independent SRWs starting at x and y respectively.

The membrane model presents several points in common, as well as challenging differences, from the more known DGFF. The former lacks some key features of the latter, namely

(a) the random walk representation for the Green's function expressed by (1.2). Since (1.5) holds it is not possible to directly relate the biharmonic operator to some function of the random walk, and this deprives us of many tools which are at disposal for the DGFF.

(b) Absence of monotonicity, for example the FKG inequality:

$$\mathbf{E}_N[fg] \geq \mathbf{E}_N[f]\mathbf{E}_N[g]$$

for any two monotonically increasing functions f and g.

1.3. High points for the membrane model in the critical dimension

This section is based on a paper which was published in *Electronic Journal of Probability*. We study the fractal structure of the set of high points for the membrane model in the critical dimension $d = 4$. The membrane model is a centered Gaussian field whose covariance is the inverse of the discrete bilaplacian operator on $\mathbb{Z}^4$. We are able to compute the Hausdorff dimension of the set of points which are atypically high, and also that of clusters, showing that high points tend not to be evenly spread on the lattice. We will see that these results follow closely those obtained by O. Daviaud [23] for the 2-dimensional discrete Gaussian Free Field.

1.3.1. Main results

The study of exceedencees of Gaussian fields is well-known and applies to the membrane model too. Despite the lack of several tools which are present for the DGFF for instance it is sufficient to establish two crucial properties to study the high points of the probability measure (1.4): one is the *logarithmic bound* on covariances, and the other one is the *2-Markov property*, which can be stated as follows:

Definition 1.7 (2-Markov property)**.** *Let $A, B \subseteq V_N$ and $\operatorname{dist}(A,B) \geq 3$. Then $\{\varphi_x\}_{x\in A}$ and $\{\varphi_x\}_{x\in B}$ are independent under the conditional law*

$$\mathbf{P}_N\left(\cdot \,|\, \sigma\left(\{\varphi_x,\, x \notin A \cup B\}\right)\right).$$

Intuitively, the 2-Markov property follows from the interaction due to the Bilaplacian matrix, whose non-vanishing terms involve up to 2nd nearest-neighbors. This means in addition that the model presents finite-range interactions. This property suggests that the behavior of certain Gaussian fields with respect to exceedences is universal, in the sense that as soon as the model displays a Gibbs-Markov property and covariances

decay at the same rate, then the behavior of high points is the same (with some small adjustments to be done according to the dimension). This also opens up the question of whether there are other points in common between log-correlated Gaussian fields, and we believe a more precise answer will be given soon.
The starting point is understanding how many "high" points viz. points that grow more than the average there are typically. The first step is to find the average height of the field, in other words to show that there exists a constant $c > 0$ such that

$$\mathbf{E}\left(\max_{x\in V_N}\varphi_x\right) / \log N \overset{N\to+\infty}{\longrightarrow} c.$$

Theorem 1.8 ([57, Theorem 1.2]). *Let $d = 4$, $\ell \in (0, 1)$,*

$$V_N^\ell := \{x \in V_N : d(x, V_N^c) \geq \ell N\} \tag{1.6}$$

and let $g := 8/\pi^2$. Then

(a)

$$\lim_{N\to+\infty} \mathbf{P}\left(\sup_{x\in V_N}\varphi_x \geq 2\sqrt{2g}\log N\right) = 0.$$

(b) If $0 < \ell < 1/2$, $0 < \eta < 1$ there exists $C = C(\ell, \eta) > 0$ such that

$$\mathbf{P}\left(\sup_{x\in V_N^\ell}\varphi_x \geq \left(2\sqrt{2g} - \eta\right)\log N\right) \leq \exp\left(-C\log^2 N\right).$$

Roughly said, the first-order approximation of the maximum is of order $\log N$, which also implies that the field behaves approximately like independent variables. For us then an *α-high point* will be a point whose height is greater than $2\sqrt{2g}\alpha\log N$. The behavior of α-high points for the 2-dimensional DGFF, as shown in [23], tells us that such points exhibit a fractal structure. Very similar results were obtained by Dembo, Peres, Rosen and Zeitouni in [27] for the set of late points of the 2-d standard random walk.
To begin with, we recall the definition of the discrete fractal dimension:

Definition 1.9 (Discrete fractal dimension, [7]). *Let $A \subseteq \mathbb{Z}^d$. If the following limit exists, the fractal dimension of A is*

$$dim(A) := \lim_{N\to\infty}\frac{\log|A\cap V_N|}{\log N}.$$

The fractal dimension of the high points is given then in

Theorem 1.10 (Number of high points)**.** *Let* $\ell \in (0,1)$, *and*

$$\mathcal{H}_N(\eta) := \left\{ x \in V_N^\ell : \varphi_x \geq 2\sqrt{2g}\eta \log N \right\}$$

be the set of η*-high points, where* $g := \frac{8}{\pi^2}$.

(a) For $0 < \eta < 1$ *we obtain the following limit in probability:*

$$\lim_{N \to +\infty} \frac{\log |\mathcal{H}_N(\eta)|}{\log N} = 4(1 - \eta^2).$$

(b) For all $\delta > 0$ *there exists a constant* $C > 0$ *such that for* N *large*

$$\mathbf{P}_N \left(\left\{ |\mathcal{H}_N(\eta)| \leq N^{4(1-\eta^2)-\delta} \right\} \right) \leq \exp(-C \log^2 N).$$

We can push further the comparison between the DGFF and the Membrane Model at their respective critical dimensions, and one can find an interesting similarity in the behavior of the points. [23] for example also showed that high points appear in clusters; this is what occurs in the membrane model, as the following two theorems show:

Theorem 1.11 (Cluster of high points 1)**.** *Let*

$$D(x, \rho) := \{ y \in V_N : |y - x| \leq \rho \}.$$

For $0 < \alpha < \beta < 1$ *and* $\delta > 0$

$$\lim_{N \to +\infty} \max_{x \in V_N^\ell} \mathbf{P}_N \left(\left| \frac{\mathcal{H}_N(\alpha) \cap D(x, N^\beta)}{\log N} - 4\beta(1 - (\alpha/\beta)^2) \right| > \delta \right) = 0. \tag{1.7}$$

Theorem 1.12 (Cluster of high points 2)**.** *For* $0 < \alpha < 1$, $0 < \beta < 1$ *and* $\delta > 0$ *we have*

$$\lim_{N \to +\infty} \max_{x \in V_N^\ell} \mathbf{P} \left(\left| \frac{\log |\mathcal{H}_N(\alpha) \cap D(x, N^\beta)|}{\log N} - 4\beta(1 - \alpha^2) \right| > \delta \,|\, x \in \mathcal{H}_N(\alpha) \right) = 0.$$

It is also possible to evaluate the average number of pairs of high points as in the following theorem:

Theorem 1.13 (Pairs of high points)**.** *Let* $0 < \alpha < 1$, $0 < \beta < 1$ *and let*

$$F_{h,\beta}(\gamma) := \gamma^2(1 - \beta) + \frac{h(1 - \gamma(1 - \beta))^2}{\beta}$$

$$\Gamma_{\alpha,\beta} := \left\{ \gamma \geq 0 : 4 - 4\beta - 4\alpha^2 F_{0,\beta}(\gamma) \geq 0 \right\} = \left\{ \gamma \geq 0 : (1 - \alpha^2\gamma^2) \geq 0 \right\},$$

$$\rho(\alpha,\beta) := 4 + 4\beta - 4\alpha^2 \inf_{\gamma \in \Gamma_{\alpha,\beta}} F_{2,\beta}(\gamma) > 0.$$

Note that $\Gamma_{\alpha,\beta} = [0,\, 1/\alpha]$ *is independent of* β*. Then the following limit in probability holds:*

$$\lim_{N\to+\infty} \frac{\log\left|\{(x,y) \in \mathcal{H}_N(\alpha) :\, |x-y| \le N^\beta\}\right|}{\log N} = \rho(\alpha,\beta).$$

Finally we can also show what the maximum width of a spike of given length is:

Theorem 1.14 (The biggest high square). *Let* $-1 < \eta < 1$, $D_N(\eta)$ *the side length of the biggest sub-box for which all height variables are uniformly greater than* $2\sqrt{2g}\eta\log N$*, i. e.*

$$D_N(\eta) := \sup\left\{a \in \mathbb{N} :\, \exists x \in V_N^\ell :\, \min_{y\in B(x,a)} \varphi_y \ge 2\sqrt{2g}\eta \log N\right\}.$$

Then the following limit in probability holds:

$$\lim_{N\to+\infty} \frac{\log D_N(\eta)}{\log N} = \frac{1-\eta}{2}.$$

The paper is organized as follows: in Section 2 we will prove some preliminary results that will be used for the proofs of the main theorems, to which Section 3 is going to be devoted.

1.4. Preliminary Lemmas and results

Notation

$D(x,a)$ (resp. $D(x,a]$) denotes the open (resp. closed) Euclidean ball of center x and radius a, while $B(x,a)$ is a box centered at x of side length a. For the rest of this notice, recall the definition (1.6) and we let once and for all $\ell \in (0,1/2)$. Let $x_0 \in V_N$ and

$$M_\alpha := \left\{x_0 + i(N^\alpha + 4) :\, i \in N^4 \text{ and } x_0 + i(N^\alpha + 2) \subset V_N\right\}.$$

We denote by x_B the center of a (sub)box B and as Π_α the union of sub-boxes of side-length N^α (without discretization issues) and midpoint in M_α. $\mathscr{F}_\alpha$ will be the sigma-algebra generated by $\{\varphi_x\}$ for $x \in \bigcup_{B\in\Pi_\alpha} \partial_2 B$. Practically we denote with Π_α a set of disjoint boxes separated by layers of thickness 2, which thanks to the 2-Markov property will enable us to perform a decomposition procedure on these sets.
Furthermore $\varphi_B := \mathbf{E}\left(\varphi_{x_B} | \mathscr{F}_{\partial_2 B}\right)$ and $\mathbf{Var}_B(\varphi_x) := \mathbf{Var}_N\left(\varphi_x | \mathscr{F}_{\partial_2 B}\right)$.

1.4.1. Lemmas

The function $\overline{G}_N(\cdot,\cdot)$

In order to prove some of the next results we will introduce the convolution of the harmonic Green's function, which will prove to be a key tool to obtain the crucial estimates on the covariances of our model. Let A be an arbitrary subset of $\mathbb{Z}^4$, and for $x \in A$ let $\Gamma_A(x,\cdot)$ be the solution of the discrete boundary value problem

$$\begin{cases} \Delta\Gamma_A(x,y) = \delta_{xy}, & y \in A \\ \Gamma_A(x,y) = 0, & y \in \partial A. \end{cases}$$

Note that Γ_N as in (1.2) is the unique solution to the above problem for $A := V_N$. The convolution of Γ_N is

$$\overline{G}_N(x,y) := \sum_{z\in V_N} \Gamma_N(x,z)\Gamma_N(z,y), \quad x,\, y \in V_N.$$

[57] contains several bounds and properties of such a function, and we would like here to recall those that we are going to use in the sequel: for all $x,\, y \in V_N$

- symmetry: $\overline{G}_N(x,y) = \overline{G}_N(y,x)$,
- [57, Lemma 2.2] if $\ell \in (0,1/2)$ there exist $c_1 = c_1(\ell) > 0$, $c_2 > 0$ such that

$$g\log N + c_1 \le \overline{G}_N(x,y) \le g\log N + c_2 \tag{1.8}$$

With this in mind it is now easier for us to show how to bound the variances and covariances of our field.

Lemma 1.15 (Bounds on the variances)**.** *Let $d = 4$ and $0 < \delta < 1$. Then*

- *there exists $C > 0$ such that*

$$\sup_{x\in V_N} \mathbf{Var}_N(\varphi_x) \le g\log N + C. \tag{1.9}$$

- *There exists $C(\ell) > 0$ such that*

$$\sup_{x\in V_N^\ell} |\mathbf{Var}_N(\varphi_x) - g\log N| \le C(\ell). \tag{1.10}$$

- *There exist $C > 0$ and $C(\ell) > 0$ such that*

$$\sup_{\substack{x,y\in V_N^\ell \\ x\neq y}} \mathbf{Cov}_N(\varphi_x,\varphi_y) - g(\log N - \log|x-y|) \le C. \tag{1.11}$$

$$\sup_{\substack{x,y\in V_N^\ell \\ x\neq y}} \left|\mathbf{Cov}_N(\varphi_x,\varphi_y) - g(\log N - \log|x-y|)\right| \le C(\ell). \tag{1.12}$$

Proof. For the variances see [57, Proposition 1.1]. For the covariances, remember that in [57, Corollary 2.9] that for all $d \geq 4$ and for all $x \in V_N^\ell$

$$\sup_{y \in V_N^\ell} |G_N(x,y) - \overline{G}_N(x,y)| \leq c = c(\ell) < +\infty. \tag{1.13}$$

It is therefore sufficient to show that (1.11) and (1.12) hold for $\overline{G}(\cdot,\cdot)$. But we have from [57, Lemma 2.10], that there exists a constant K such that in $d = 4$ for $x \neq y$ and all $\alpha \in (0,2)$

$$\overline{G}_N(x,x) - \overline{G}_N(x,y) = g \log |y-x| + K + o\left(|y-x|^{-\alpha}\right).$$

Hence

$$\begin{aligned} G_N(x, y) &\leq \overline{G}_N(x, y) + c = \overline{G}_N(x, x) - g \log |y-x| + K' \overset{(1.8)}{\leq} \\ &\leq g \log N - g \log |y-x| + K'. \end{aligned}$$

The other bound follows similarly by considering (1.12). □

Next we give a decomposition of the field which is similar to the one existing for the DGFF (see for example [77, Section 2.1]). With this in mind, we can prove that conditioning on the values of the field assumed on the double boundary of a subset of $V_N \subseteq \mathbb{Z}^4$ (in fact of any $\mathbb{Z}^d$) the resulting field is again the membrane model restricted to the interior of the smaller domain.

Lemma 1.16. *Let $B \subseteq V_N$. Let $\mathscr{F} := \sigma(\varphi_z,\, z \in V_N \setminus B)$. Then*

$$\{\varphi_x\}_{x \in B} \overset{d}{=} \{\mathbf{E}_N[\varphi_x | \mathscr{F}] + \psi_x\}_{x \in B}$$

where "$\overset{d}{=}$" indicates equality in distribution, in particular under $\mathbf{P}_N(\cdot)$

(a) $\psi_x \perp\!\!\!\perp \mathscr{F}$*;*

(b) $\{\psi_x\}_{x \in B}$ *is distributed as the membrane model with 0-boundary conditions on B.*

Proof. Set $\psi_x := \varphi_x - \mathbf{E}[\varphi_x | \mathscr{F}]$ for all $x \in B$. We have to show that the above results hold.

(a) It is clear from the definition.

(b) Being $\mathbf{P}_N$ a Gibbs measure, it satisfies the DLR equation: for all $A \subseteq V_N$, $\mathscr{F}_{A^c} := \sigma(\varphi_z,\, z \in A^c)$,

$$\mathbf{P}_N(\cdot \mid \mathscr{F}_{A^c})(\eta) = P_{A,\eta}(\cdot) \quad \mathbf{P}_N(\mathrm{d}\eta) - a.s. \tag{1.14}$$

with

$$P_{A,\eta}(\mathrm{d}\varphi) = \frac{1}{Z_A} \exp\left(-\frac{1}{2} \sum_{x \in \mathbb{Z}^d} (\Delta\varphi_x)^2\right) \prod_{x \in A} \mathrm{d}\varphi_x \prod_{x \in V_N \setminus A} \delta_{\eta_x}(\mathrm{d}\varphi_x).$$

In other words, $P_{A,\eta}$ is a Gaussian distribution with covariance matrix $\left(\Delta_A^2\right)^{-1}$. Since $\mathbf{Cov}_N(\cdot,\cdot|\mathscr{F}_{A^c})$ we find out that it equals G_A. In our case this means that $\mathbf{Cov}_N(\cdot|\mathscr{F})$ is deterministic and equal to G_B. So

$$\mathbf{Cov}_N(\psi_x,\,\psi_y) = \mathbf{Cov}_N\left(\psi_x,\,\psi_y|\mathscr{F}\right) = \mathbf{Cov}_N\left(\varphi_x,\,\varphi_y|\mathscr{F}\right) = G_B(x,y)$$

□

Remark 1.17. *This result gives us a decomposition of the membrane model in all dimensions.*

Lemma 1.18. *Let* $0<\alpha<1$ *and* $0<\beta<1$, $\delta>0$ *and we define*

$$S = S(\epsilon) := \left\{(x,y)\in V_N^\ell:\, N^{\beta(1-\epsilon)} \le |x-y| \le N^\beta\right\}.$$

Then there exist C, $\epsilon_0>0$ *(which can be chosen uniformly on* (α,β) *on compact sets of* $(0,1)^4$*) and* $\gamma_\star := 2(2-\beta)^{-1}$ *such that for all* $\epsilon\le\epsilon_0$ *and all* N

$$\max_{(x,y)\in S}\mathbf{P}\left(x,y\in\mathcal{H}_N(\alpha)\right)\le CN^{-4\alpha^2F_{2,\beta}(\gamma_\star)+\delta}.$$

Proof. Let $Z:=\varphi_x+\varphi_y$ and we see that

$$\{x,y\in\mathcal{H}_N(\alpha)\}\subseteq\left\{Z\ge 4\sqrt{2g}\alpha\log N\right\}.$$

We obtain also from (1.11) that

$$\mathbf{Cov}_N(\varphi_x,\varphi_y)\le g\log N - g\beta(1-\epsilon)\log N + O(1).$$

Thus by (1.13) and (1.9)

$$\mathbf{Var}_N(Z)\le(2g(2-\beta)+O(\epsilon)+O(1/\log N))\log N.$$

Since $F_{2,\beta}(\gamma_\star)=\gamma_\star$, using (1.15)

$$\begin{aligned}
&\mathbf{P}(Z\ge 4\sqrt{2g}\alpha\log N)\le\\
\le\;&\exp\left(-\frac{16(\sqrt{2g})^2\alpha^2\log^2 N}{2((2g(2-\beta)+O(\epsilon)+O(1/\log N))\log N}\right)\le\\
\le\;&\exp\left(-4\alpha^2\gamma^*(1+O(\epsilon)+O(1/\log N))\log N\right)\le\\
\le\;&CN^{-4\alpha^2F_{2,\beta}(\gamma_\star)+O(\epsilon)}.
\end{aligned}$$

□

Lemma 1.19. *Let* $B := B(x, 4N^\beta)$, $\epsilon > 0$, $b^\pm(\alpha,\beta,\epsilon,N) = 2\sqrt{2g}(\alpha(1-\beta)\pm\epsilon)\log N$, $I(\alpha,\beta,\epsilon,N) := [b^-(\alpha,\beta,\epsilon,N), b^+(\alpha,\beta,\epsilon,N)]$. *Then*

$$\max_{x\in V_N^\ell} \mathbf{P}(\varphi_B \notin I(\alpha,\beta,\epsilon,N)) | \varphi_x \geq 2\sqrt{2g}\alpha \log N) \overset{N\to+\infty}{\longrightarrow} 0.$$

Proof. We shorten I, b^+ and b^- for the above quantities. We recall here two useful facts about normal random variables (whose short proof is postponed to the appendix). If $X \sim \mathcal{N}(0,1)$ then

$$P(|X| \geq a) \leq \exp(-a^2/2), \quad \forall\, a \geq 0, \tag{1.15}$$

$$P(|X| \geq a) \geq \frac{\exp(-a^2/2)}{\sqrt{2\pi}a}, \quad \forall\, a \geq 1. \tag{1.16}$$

For $\eta > 0$ we obtain with (1.15) and (1.16)

$$\mathbf{P}(\varphi_x \geq 2\sqrt{2g}\alpha(1+\eta)\log N | \varphi_x \geq 2\sqrt{2g}\alpha \log N) \to 0.$$

as $N \to +\infty$. This yields

$$\begin{aligned}
&\mathbf{P}(\varphi_B \notin I | \varphi_x \geq 2\sqrt{2g}\alpha \log N) = o(1) + \\
&+\mathbf{P}(\varphi_B \notin I, \varphi_x \leq 2\sqrt{2g}\alpha(1+\eta)\log N | \varphi_x \geq 2\sqrt{2g}\alpha \log N) \leq \\
&\leq o(1) + \mathbf{P}(\varphi_B \notin I | \varphi_x \in (1, 1+\eta) 2\sqrt{2g}\alpha \log N).
\end{aligned}$$

Now we write $\varphi_x = \varphi_x - \varphi_B + \varphi_B$ and observe that $\varphi_B \perp\!\!\!\perp \varphi_x - \varphi_B$. Therefore $\mathbf{Cov}_N(\varphi_x, \varphi_B) = \mathbf{Var}_N(\varphi_B)$ and so there exists $Z \sim \mathcal{N}(0, \sigma_Z^2)$, $\sigma_Z^2 > 0$, for which

$$\varphi_B = \frac{\mathbf{Var}_N(\varphi_B)}{\mathbf{Var}_N(\varphi_x)}\varphi_x + Z, \quad Z \perp\!\!\!\perp \varphi_x.$$

If x is the center of $B \subseteq C$ we can decompose the variances as $\mathbf{Var}_C(\varphi_x) = \mathbf{Var}_C(\varphi_B) + \mathbf{Var}_B(\varphi_x)$, and with this

$$\frac{\mathbf{Var}_N(\varphi_B)}{\mathbf{Var}_N(\varphi_x)} = (1-\beta) + O\left(\frac{1}{\log N}\right).$$

It must then be that $\mathbf{Var}_N(Z) = O(\log N)$. Consequently

$$\begin{aligned}
&\mathbf{P}(\varphi_B \geq b^+ | \varphi_x \in (1, 1+\eta) 2\sqrt{2g}\alpha \log N) \leq \\
&\leq \mathbf{P}\left(Z + \left((1-\beta) + O\left(\frac{1}{\log N}\right)\right)(1+\eta)2\sqrt{2g}\alpha \log N \geq b^+\right) \to 0
\end{aligned}$$

for $\eta < \epsilon/(\alpha(1-\beta))$. Similarly

$$\begin{aligned}
&\mathbf{P}(\varphi_B \leq b^- | \varphi_x \in (1, 1+\eta) 2\sqrt{2g}\alpha \log N) \leq \\
&\leq \mathbf{P}\left(Z + \left((1-\beta) + O\left(\frac{1}{\log N}\right)\right)2\sqrt{2g}\alpha \log N \leq b^-\right) \to 0.
\end{aligned}$$

□

Lemma 1.20. *We keep the notation of Lemma 1.18. Let $0 < \alpha < \beta < 1$ and $\delta > 0$. For $(x, y) \in S$ define $T(x, y)$ as the set of sub-boxes of side length $2N^{\beta}$ such that the centered subbox of side length N^{β} contains x, y. Then we can find $C, \epsilon_0 > 0$ such that for $\epsilon \leq \epsilon_0$ and all N*

$$\begin{aligned}&\max_{\substack{x,y\in S\\ B\in T(x,y)}} \mathbf{P}\left(\{x, y \in \mathcal{H}_N(\alpha)\} \cap \left\{\varphi_B \leq 2\sqrt{2g}\alpha\gamma(1-\beta)\log N\right\}\right)\\ &\leq CN^{-4\alpha^2 F_{2,\beta}(\min\{\gamma,\gamma_\star\})+\delta}.\end{aligned}$$

ϵ_0 can be chosen uniformly on (α, β) on compact sets of $(0,1)^4$.

Proof. Define

$$E := \{x, y \in \mathcal{H}_N(\alpha)\} \cap \left\{\varphi_B \leq 2\sqrt{2g}\alpha\gamma(1-\beta)\log N\right\}.$$

We distinguish two cases:

$\gamma \geq \gamma_\star$. We have $\mathbf{P}(E) \leq \mathbf{P}(\{x, y \in \mathcal{H}_N(\alpha)\})$: the claim follows from Lemma 1.18 because $\min\{\gamma, \gamma_\star\} = \gamma_\star$.

$\gamma < \gamma_\star$. It follows from the definition of $\gamma_\star$ that $\gamma < \gamma_\star$ implies $\gamma < 2(2-\beta)^{-1}$. For this reason set $a := 1 - \gamma(1-\beta) > 0$ and $b := \gamma(2-\beta) - 2 < 0$. Letting $Z := a(\varphi_x + \varphi_y) + b\varphi_B$

$$E \subseteq \left\{Z \geq (2a + b\gamma(1-\beta))\alpha 2\sqrt{2g}\log N\right\}.$$

Furthermore we have the usual decomposition

$$\begin{aligned}\mathbf{Var}_N(Z) &= a^2\mathbf{Var}_N(\varphi_x) + a^2\mathbf{Var}_N(\varphi_y) + b^2\mathbf{Var}_N(\varphi_B) +\\ &+2ab\mathbf{Cov}_N(\varphi_x, \varphi_B) + 2ab\mathbf{Cov}_N(\varphi_y, \varphi_B) +\\ &+2a^2\mathbf{Cov}_N(\varphi_x, \varphi_y).\end{aligned} \tag{1.17}$$

By Lemma 1.15

$$\mathbf{Var}_N(\varphi_B) = \mathbf{Var}_N(\varphi_{x_B}) - \mathbf{Var}(\varphi_{x_B}|\mathscr{F}_{\partial_2 B}) \leq g(1-\beta)\log N + O(1).$$

and

$$\begin{aligned}\mathbf{Cov}_N(\varphi_x, \varphi_B) &= \mathbf{E}(\mathbf{E}(\varphi_x|\mathscr{F}_{\partial_2 B})\mathbf{E}(\varphi_{x_B}|\mathscr{F}_{\partial_2 B})) =\\ &= \mathbf{Cov}_N(\varphi_x, \varphi_{x_B}) - \mathbf{Cov}(\varphi_x, \varphi_{x_B}|\mathscr{F}_{\partial_2 B}) \geq\\ &\geq g(\log N - \log|x - x_B|) - g(\beta\log N - \log|x - x_B|) + O(1) =\\ &= g(1-\beta)\log N + O(1).\end{aligned}$$

Analogously

$$\mathbf{Cov}_N(\varphi_y, \varphi_B) \geq g(1-\beta)\log N + O(1).$$

Define the auxiliary function $f(a,b,\beta) := 2a^2(2-\beta) + b^2(1-\beta) + 4ab(1-\beta)$. We use these bounds in (1.17) to obtain

$$\mathbf{Var}_N(Z) \leq (f(a,b,\beta) + O(\epsilon) + O(1/\log N))g\log N.$$

By the equality $2a + b = \gamma\beta$

$$4a^2 + b^2 + 4ab = (2a+b)^2 = \gamma^2\beta^2.$$

Then

$$\begin{aligned} f(a,b,\beta) &= (2a+b)^2 - \beta(2a^2 + b^2 + 4ab) = \\ &= (4a^2 + b^2 + 4ab)(1-\beta) + 2\beta a^2 = \\ &= (\gamma\beta)^2(1-\beta) + 2\beta a^2 = \\ &= \beta(\beta\gamma^2(1-\beta) + 2a^2) = \\ &= \beta((2a+b)(1-a) + 2a^2) = \\ &= \beta(2a + b - ab). \end{aligned}$$

Hence

$$\mathbf{Var}_N(Z) \leq (\beta(2a+b-ab) + O(\epsilon) + O(1/\log N))g\log N. \tag{1.18}$$

Since $2a + b - ab = 2a + b\gamma(1-\beta)$ (1.17) and (1.18) yield

$$\mathbf{P}(E) \leq C\exp\left(-\left(\frac{4\alpha^2(2a+b-ab)}{\beta} + O(\epsilon)\right)\log N\right).$$

Finally notice that

$$\begin{aligned} &\beta F_{2,\beta}(\gamma) = \beta\gamma^2(1-\beta) + 2(1-\gamma(1-\beta))^2 = \beta\gamma^2(1-\beta) + 2a^2 = \\ &= (2a+b)(1-a) + 2a^2 = 2a + b - ab. \end{aligned}$$

This allows us to conclude the proof. □

Finally we would like to recall

Lemma 1.21 ([57, Lemma 2.11]). *Let $0 < n < N$, $A_N \subseteq \mathbb{Z}^4$ be a box of side-length N, $A_n \subseteq A_N$ a box of side-length n. Let $0 < \epsilon < 1/2$. There exists $C > 0$ such that for all $x \in A_n$ with $|x - x_B| < \epsilon n$*

$$\mathbf{Var_N}\left(\mathbf{E}\left(\varphi_x \mid \mathscr{F}_{\partial_2 A_n}\right) - \mathbf{E}_N\left(\varphi_{x_B} \mid \mathscr{F}_{\partial_2 A_n}\right) \mid \mathscr{F}_{\partial_2 A_N}\right) \leq C\epsilon.$$

Remark 1.22. *In [57] the above Lemma is stated with the assumption that "*the boxes A_n and A_N have the same center*". However one sees that the result can be obtained removing this condition which is not necessary.*

1.5. Five theorems

Proof of Theorem 1.10. The core of the proof is the lower bound (b) which was already proved by [57, Theorem 1.3] and is based on the hierarchical decomposition of the membrane model, similar to that of the DGFF (for the main idea supporting the proof we also refer to [14]). We show here for the reader's convenience the upper bound, in order to obtain the desired limit in probability.

Proof of Theorem 1.10 (a). For any $\delta > 0$ one can apply Chebyshev's inequality to get

$$
\begin{aligned}
&\mathbf{P}\left(\left\{|\mathcal{H}_N(\eta)| \le N^{-4(1-\eta^2)-\delta}\right\}\right) \le N^{4(1-\eta^2)+\delta}\mathbf{E}|\mathcal{H}_N(\eta)| \le \\
\le\ & N^{-4(1-\eta^2)-\delta} N^4 \max_{x\in V_N} \mathbf{P}\left(\varphi_x \ge 2\sqrt{2g}\eta \log N\right) \le \\
\le\ & N^{-4(1-\eta^2)-\delta} N^4 \exp\left(-\frac{8g\eta^2 \log^2 N}{2g\log N + C}\right) \le N^{-4(1-\eta^2)-\delta} N^{4-4\eta^2} \to 0
\end{aligned}
$$

where we have used Lemma 1.15 too.

Proof of Theorem 1.11. We choose $\eta, \delta > 0$ and define

$$
D_+ := \left\{\varphi_B \le 2\sqrt{2g}\eta \log N\right\},
$$

$$
C_+ := \left\{|\mathcal{H}_N(\alpha) \cap D(x, N^\beta)| \ge N^{4\beta(1-(\alpha/\beta)^2)-\delta)}\right\}
$$

and for an $\epsilon > 0$ to be fixed later

$$
A := \bigcup_{y \in B(x,N^\beta)} \left\{|\mathbf{E}(\varphi_y|\mathscr{F}_{\partial_2 B}) - \varphi_B| \ge 2\sqrt{2g}\epsilon \log N\right\}.
$$

By Lemma 1.21 $\mathbf{Var}_N(\varphi_B - \mathbf{E}(\varphi_y|\mathscr{F}_{\partial_2 B})) \le c$ (we may assume that $B(x,N^\beta) \subsetneq V_N^\ell$), and so

$$
\mathbf{P}(A) = O\left(N^{4\beta}\exp\left(-c\log^2 N\right)\right)
$$

tends to 0. Furthemore also $\mathbf{P}(D_+^c)$ tends to 0 by virtue of the bounds on covariances and (1.15). We then have

$$
\begin{aligned}
&\mathbf{P}(C_+) = \mathbf{E}(\mathbf{P}(C_+|\mathscr{F}_{\partial_2 B})) \le \mathbf{P}(A) + \mathbf{P}(D_+^c) + \mathbf{E}(\mathbf{P}(C_+|\mathscr{F}_{\partial_2 B})\mathbf{1}_{A^c\cap D_+}) \le \\
&\le o(1) + \mathbf{P}\left(\left|\mathcal{H}_{4N^\beta}\left(\frac{\alpha-\epsilon'}{\beta}\right)\right| \ge N^{4\beta(1-(\alpha/\beta)^2)-\delta)}\right)
\end{aligned}
$$

where ϵ' satisfies

$$
\frac{\alpha - \epsilon'}{\beta}\log(4N^\beta) = (\alpha - \eta - \epsilon)\log N.
$$

By tuning the parameters N large enough and η, ϵ small enough we can obtain

$$4\beta\left(1-\left(\frac{\alpha-\epsilon'}{\beta}\right)^2\right) < 4\beta\left(1-\left(\frac{\alpha}{\beta}\right)^2\right)+\delta$$

(roughly speaking, we have $\epsilon' \approx \alpha(1-\beta)$). By Theorem 1.10

$$\mathbf{P}\left(\left|\mathcal{H}_{4N^\beta}\left(\frac{\alpha-\epsilon'}{\beta}\right)\right| \geq N^{4\beta(1-(\alpha/\beta)^2)-\delta}\right) \to 0$$

and from this the claim follows. We now go to the lower bound proof, which is similar in spirit to the upper bound. By setting

$$D_- := \left\{\varphi_B \geq -2\sqrt{2g}\eta \log N\right\},$$

$$C_- := \left\{|\mathcal{H}_N(\alpha) \cap D(x, N^\beta)| \leq N^{4\beta(1-(\alpha/\beta)^2)-\delta)}\right\}$$

we also define

$$\mathcal{H}_N^s(\eta) := \left\{x \in V_N^s : \varphi_x \geq 2\sqrt{2g}\eta \log N\right\}, \quad s \in (0, 1/2).$$

We observe that

$$\begin{aligned}
&\mathbf{P}(C_-) = \mathbf{E}(\mathbf{P}(C_-|\mathscr{F}_{\partial_2 B})) \leq \mathbf{P}(A) + \mathbf{P}(D_-^c) + \mathbf{E}(\mathbf{P}(C_+|\mathscr{F}_{\partial_2 B})\mathbf{1}_{A^c \cap D_-}) \leq \\
&\leq o(1) + \mathbf{P}\left(\left|\mathcal{H}_{4N^\beta}^{3/8}\left(\frac{\alpha+\epsilon'}{\beta}\right)\right| \leq N^{4\beta(1-(\alpha/\beta)^2)-\delta)}\right)
\end{aligned}$$

where ϵ' satisfies

$$\frac{\alpha+\epsilon'}{\beta}\log(4N^\beta) = (\alpha+\eta+\epsilon)\log N$$

and we conclude as before.

Proof of Theorem 1.12. We will use the notation $b^{\pm}(\alpha, \beta, \eta, N)$ as in the proof of Lemma 1.19. We will also introduce the following quantities: let $B := B(x, 4N^\beta)$, and for $\eta, \delta > 0$,

$$\begin{aligned}
E &:= \left\{|\mathcal{H}_N(\alpha) \cap D(x, N^\beta)| \leq N^{4\beta(1-\alpha^2)-\delta}\right\}, \\
F &:= \left\{\varphi_B \geq b^-(\alpha, \beta, \eta, N)\right\}, \\
G &:= \left\{x \in \mathcal{H}_N(\alpha)\right\}.
\end{aligned}$$

Lower bound. Thanks to the proof of Lemma 1.19 we have $\mathbf{P}(E|G) = \mathbf{P}(E|F \cap G)\mathbf{P}(F|G) + o(1) = \mathbf{P}(E|F \cap G)(1 + o(1)) + o(1)$. This means that

$$\begin{aligned}
\mathbf{P}(E|F,G) &= \frac{\mathbf{P}(E \cap F \cap G)}{\mathbf{P}(F \cap G)} \leq \frac{1}{\mathbf{P}(F \cap G)}\sqrt{\mathbf{P}(G)\mathbf{P}(E \cap F)} = \\
&= \frac{1}{\mathbf{P}(F|G)\mathbf{P}(G)}\sqrt{\mathbf{P}(G)\mathbf{P}(F)\mathbf{P}(E|F)} = \\
&\overset{Lemma\,1.19}{=} (1+o(1))\sqrt{\frac{\mathbf{P}(F)}{\mathbf{P}(G)}\mathbf{P}(E|F)}.
\end{aligned}$$

We know by the bounds (1.9) and (1.16)

$$\begin{aligned}
\mathbf{P}(G) &= \mathbf{P}(\varphi_x \geq 2\sqrt{2g}\alpha \log N) \geq c_1 \frac{\exp\left(-\frac{8g\alpha^2 \log^2 N}{2g \log N + c_2}\right)}{c_3 \log N} \geq \\
&\geq \exp\left(-d' \log N\right), \\
\mathbf{P}(F) &= \mathbf{P}(\varphi_B \geq 2\sqrt{2g}(\alpha(1-\beta) - \eta) \log N) \leq \\
&\leq c_4 \exp\left(-\frac{8g(\alpha(1-\beta)-\eta)^2 \log^2 N}{2g(1-\beta)\log N + c_5}\right) \leq \exp\left(-d'' \log N\right)
\end{aligned}$$

for some $d', d'' > 0$. Therefore we can find $d > 0$ such that $\mathbf{P}(F)/\mathbf{P}(G) \leq \exp(d \log N)$ and to show the result it suffices to prove that $\mathbf{P}(E|F) \leq \exp(-c \log^2 N)$ for a positive c. For this purpose define

$$A := \bigcup_{y \in B} \left\{ |\mathbf{E}(\varphi_y | \mathscr{F}_{\partial_2 B}) - \varphi_B| \geq 2\sqrt{2g}\epsilon \log N \right\}.$$

From Lemma 1.21 it follows that $\mathbf{P}(A) \leq \exp(-c \log^2 N)$ for $c > 0$ and from (1.16) that $\mathbf{P}(F) \geq \exp\left(-d \log N\right)$ for some $d > 0$, all in all $\mathbf{P}(A|F) \leq \exp\left(-O\left(\log^2 N\right)\right)$. So we can write

$$\begin{aligned}
&\mathbf{P}(E|F) \leq \frac{\mathbf{P}(F \cap A)}{\mathbf{P}(F)} + \\
&+ \frac{\mathbf{P}(E \cap F \cap A^c)}{\mathbf{P}(F)} \leq \\
&\exp\left(-O\left(\log^2 N\right)\right) + \frac{\mathbf{E}(\mathbf{P}\left(E|\mathscr{F}_{\partial_2 B}\right)\mathbf{1}_{A^c}\mathbf{1}_F)}{\mathbf{P}(F)}.
\end{aligned}$$

If we are on $A^c \cap F$, then

$$\begin{aligned}
&\mathbf{P}\left(|\mathcal{H}_N(\alpha) \cap D(x, N^\beta)| \leq N^{4\beta(1-\alpha^2)-\delta} | \mathscr{F}_{\partial_2 B}\right) \leq \\
&\leq \mathbf{P}\left(\left|\mathcal{H}^{3/8}_{4N^\beta}(\alpha + \epsilon')\right| \leq N^{4\beta(1-\alpha^2)-\delta}\right) \qquad (1.19)
\end{aligned}$$

where ϵ' is such that

$$(\alpha - (\alpha(1-\beta) - \eta) + \epsilon) \log N = (\alpha + \epsilon') \log 4N^{\beta}. \tag{1.20}$$

From Theorem 1.10 we know that (1.19) is bounded from above by $\exp(-c \log^2 N)$ for a constant $c > 0$, provided that ϵ' is small (which can be obtained if η, ϵ and N are small, small and large respectively).

Upper bound. Let $K \in \mathbb{N}$ and $\left\{\beta_j := \frac{j}{K}\beta\right\}_{1 \le j \le K}$. Then let

$$D_1 := D\left(x, N^{\beta_1}\right), \quad D_i := D\left(x, N^{\beta_i}\right) \setminus D\left(x, N^{\beta_{i-1}}\right).$$

Since $D\left(x, N^{\beta}\right) = \cup_{1 \le i \le N} D_i$

$$\begin{aligned}&\left\{\left|\mathcal{H}_N(\alpha) \cap D\left(x, N^{\beta}\right)\right| \ge N^{4\beta(1-\alpha^2)+\epsilon}\right\} \subseteq \\ &\subseteq \bigcup_{0 \le i \le N} \left\{|\mathcal{H}_N(\alpha) \cap D_i| \ge N^{4\beta_i(1-\alpha^2)+\epsilon/2}\right\}\end{aligned}$$

as soon as N is large. It is then sufficient to prove that for all i

$$\mathbf{P}\left(|\mathcal{H}_N(\alpha) \cap D_i| \ge N^{4\beta_i(1-\alpha^2)+\epsilon/2} \,|x \in \mathcal{H}_N(\alpha)\right) \overset{N \to +\infty}{\longrightarrow} 0.$$

We can consider β_j's for which $4\beta_j(1-\alpha^2) + \epsilon/2 \le 4\beta_j$. Let $B_j := B\left(x, 4N^{\beta_j}\right)$,

$$C := \left\{|\mathcal{H}_N(\alpha) \cap D_j| \ge N^{4\beta_j(1-\alpha^2)+\epsilon/2}\right\}$$

and $b^+(\alpha, \beta_j, \eta, N)$ as above. By Lemma 1.19 we obtain

$$\mathbf{P}\left(C|x \in \mathcal{H}_N(\alpha)\right) = \mathbf{P}(C \cap \left\{\varphi_{B_j} \le b^+(\alpha, \beta_j, \eta, N)\right\} |x \in \mathcal{H}_N(\alpha)\Big) + o(1).$$

If we set $F := \left\{\varphi_{B_j} \le b^+(\alpha, \beta_j, \eta, N)\right\}$, $G := \{x \in \mathcal{H}_N(\alpha)\}$ we obtain

$$\begin{aligned}\mathbf{P}(C \cap F|G) &\overset{\text{Chebyshev inq.}}{\le} \frac{N^{-4\beta_j(1-\alpha^2)-\epsilon/2}}{\mathbf{P}(G)} \mathbf{E}(\mathbf{1}_{F\cap G}|\mathcal{H}_N(\alpha) \cap D_j|) = \\ &= \frac{N^{-4\beta_j(1-\alpha^2)-\epsilon/2}}{\mathbf{P}(G)} \mathbf{E}\left(\sum_{y \in D_j} \mathbb{1}_{\{x,y \in \mathcal{H}_N(\alpha)\}} \mathbf{1}_F\right) \le \\ &\le \frac{N^{4\beta_j\alpha^2-\epsilon/2}}{\mathbf{P}(G)} \sup_{y \in D_j} \mathbf{P}(\{x, y \in \mathcal{H}_N(\alpha)\} \cap F).\end{aligned} \tag{1.21}$$

By the bounds on the covariance and the normal distribution we have

$$\mathbf{P}(G)^{-1} \leq N^{4\alpha^2+\epsilon/8} \tag{1.22}$$

for N large. By Lemma 1.20 by defining $\gamma^* = \frac{2}{2-\beta_j} > 1$ when η is small and K large we obtain

$$\sup_{y\in D_j} \mathbf{P}(\{x,y \in \mathcal{H}_N(\alpha)\} \cap F) \leq N^{-4\alpha^2 F_{2,\beta_j}(1)+\epsilon/8} = N^{-4\alpha^2(1+\beta_j)+\epsilon/8}. \tag{1.23}$$

Inserting (1.22) and (1.23) in (1.21) we obtain

$$\mathbf{P}(C \cap F|G) \leq N^{4\beta_j\alpha^2-\epsilon/2+\epsilon/8+4\alpha^2-4\alpha^2(1+\beta_j)+\epsilon/8} = \frac{1}{N^{\epsilon/4}} \overset{N\to+\infty}{\longrightarrow} 0$$

Proof of Theorem 1.13. Preliminary we would like to make some considerations. It holds that $\rho(\alpha,\beta)$ is positive and in particular

$$\rho(\alpha,\beta) \geq 4 + 4\beta - 4\alpha^2 F_{2,\beta}(1) = 4(1-\alpha^2)(1+\beta). \tag{1.24}$$

(1.24) derives from the fact that $F_{2,\beta}(\gamma)$ has a unique global minimum at 1 in the range $\gamma \in \Gamma_{\alpha,\beta}$. Moreover notice that $\rho(\alpha,\beta)$ is increasing in β. If we set $\gamma_m := \inf_{\gamma\in\Gamma_{\alpha,\beta}} F_{2,\beta}(\gamma)$, $\gamma_* := \inf_{\gamma\geq 0}$ and $\gamma_+ := \sup\Gamma_{\alpha,\beta}$ we have $\gamma_* \leq \gamma_m \leq \gamma_+$ and moreover since $F_{h,\beta}(\cdot)$ does not depend on α as well as $\Gamma_{\alpha,\beta}$ does not depend on β we have $\gamma_m = \min\{\gamma_*,\ \gamma_+\}$. that

$$\gamma_+ = 1/\alpha \geq 1.$$

We are now ready to prove the lower and upper bounds.

Lower bound. We set

$$C := \left\{ \left| \left\{ (x,y) \in \mathcal{H}_N(\alpha) : \ |x-y| \leq N^\beta \right\} \right| \leq N^{\rho(\alpha,\beta)-\delta} \right\}.$$

Set $m_\gamma := 4 - 4\beta - 4\alpha^2 F_{0,\beta}(\gamma) = 4(1-\beta)(1-\alpha^2\gamma^2)$ and choose $\gamma < \gamma_+$ (in order to have m_γ strictly positive). Further

$$F := \left\{ B \in \Pi_\beta : \ \varphi_B \geq 2\sqrt{2g}\gamma(1-\beta)\alpha\log N \right\},$$

$$D := \left\{ |F| \geq N^{m_\gamma-\delta/2} \right\}.$$

Theorem 1.10[1] shows that $\mathbf{P}(D^c) \to 0$. Hence we rewrite

$$\mathbf{P}(C) = o(1) + \mathbf{P}(D\cap C).$$

[1]The idea is to scale the square: now we take the box with mesh N/N^β and the grid is made by $\{x_B : B \in \Pi_\beta\}$. In this way Theorem 1.10 tells us that $\mathcal{H}_{N^{1-\beta}}(\gamma\alpha) \approx N^{4(1-\beta)(1-\gamma^2\alpha^2)} = N^{m_\gamma}$.

On D we have at least $\left\{B_j:\ 1 \le j \le N^{m_\gamma - \delta/2}\right\}$ boxes. Set

$$D_j := \left\{\varphi_{B_j} \ge 2\sqrt{2g}\alpha\gamma(1-\beta)\log N\right\}.$$

We observe

$$C \cap D \subseteq E := \bigcup_{j=1}^{N^{m_\gamma-\delta/2}} \left(D_j \cap \left\{\left|\mathcal{H}_N(\alpha)\cap B_j\right| \le N^{(\rho(\alpha,\beta)-m_\gamma)/4-\delta/8}\right\}\right).$$

Let us now put for some arbitrary $\eta > 0$

$$A := \bigcup_{B\in\Pi_\beta}\ \bigcup_{y\in B\left(x_B, N^\beta/2\right)} \left\{\left|\mathbf{E}\left(\varphi_y|\mathscr{F}_B\right) - \varphi_B\right| \ge 2\sqrt{2g}\eta\log N\right\}.$$

As before $\mathbf{P}(A) = o(1)$ as $N \to +\infty$. Plugging this in, exactly as in the proof of Theorem 1.11

$$\begin{aligned}&\mathbf{P}(C\cap D) \le o(1) + \mathbf{P}(E\cap A^c) \le \\ &\le o(1) + \\ &+N^{m_\gamma-\delta/2}\mathbf{P}\left(\left|\mathcal{H}^{1/4}_{N^\beta}\left(\frac{\alpha(1-\gamma(1-\beta))+\eta}{\beta}\right)\right| \le N^{\frac{\rho(\alpha,\beta)-m_\gamma}{4}-\frac{\delta}{8}}\right).\end{aligned}$$

Finally we observe that

$$\frac{\rho(\alpha,\beta)-m_\gamma}{4} \ge 2\beta\left(1-\frac{\alpha^2\left(1-\gamma(1-\beta)\right)^2}{\beta^2}\right)$$

which is $\exp(-O(\log^2 N))$ by Theorem 1.10 for η small enough, as we have already seen. Hence $\mathbf{P}(C\cap D) = o(1)$, and we conclude the proof.

Upper bound. By Theorem 1.10 we see that for $\lambda > 0$ the number of α-high points within distance $N^{\lambda\beta}$ is at most $N^{4(1-\alpha^2)+4\lambda\beta}$. We have with (1.24) that $4(1-\alpha^2) + 4\lambda\beta \le \rho(\alpha,\beta)$ if

$$4(1-\alpha^2)+4\lambda\beta \le 4(1-\alpha^2)(1+\beta) \iff \lambda \le (1-\alpha^2).$$

Therefore when this condition is not satisfied it is enough to find that there exists $h = h(\delta) < 1$ such that for all $\beta' \in [\beta(1-\alpha^2), \beta]$

$$\mathbf{P}\left(\left|\left\{(x,y)\in\mathcal{H}_N(\alpha):\ N^{\beta'} \le |x-y| \le N^{\beta' h}\right\}\right| \ge N^{\rho(\alpha,\beta')+\delta}\right) \to 0.$$

We separate the two cases $\gamma_* \gtreqqless \gamma_m$:

$\gamma_* = \gamma_m$. Define

$$E := \left\{ \left| (x,y) \in \mathcal{H}_N(\alpha) : N^{\beta'} \le |x-y| \le N^{\beta' h} \right| \ge N^{\rho(\alpha,\beta')+\delta} \right\}.$$

By Chebyshev inequality

$$\mathbf{P}(E) \le N^{-\rho(\alpha,\beta')-\delta} \mathbf{E} \left(\sum_{(x,y):\, N^{\beta'} \le |x-y| \le N^{\beta' h}} \right) \mathbb{1}_{\{x,y \in \mathcal{H}_N(\alpha)\}} \le$$
$$\le N^{-\rho(\alpha,\beta')-\delta} N^{4+4\beta'-4\alpha^2 F_{2,\beta'}(\gamma_*)+\delta/2},$$

where we have used the assumption that h is close to 1 and Lemma 1.18

$\gamma_* > \gamma_m$. We construct for each $B \in \Pi_{\beta'}$ a bigger box of size $4N^{\beta'}$ by juxtaposing to it the 12 adjacent subboxes of same side length. We call the set of such bigger boxes $\mathcal{B}$, and for each $B' \in \mathcal{B}$ we center in $x_{B'}$ a box of twice bigger volume as B'. The latter boxes belong to a new set named $\mathcal{C}$. We remark that all pairs of points within distance $N^{\beta'}$ must belong to at least one $B' \in \mathcal{B}$. For $\epsilon > 0$ set

$$D := \left\{ \max_{C \in \mathcal{C}} \varphi_C \ge (1+\alpha\epsilon)(1-\beta')2\sqrt{2g}\log N \right\}.$$

By Lemma 1.15 and the fact that $\{\varphi_y : y \in B\}$ with boundary conditions $\partial_2 B$ is a Gaussian field

$$\mathbf{P}(D^c) \quad \le \quad \left|\Pi_{\beta'}\right| \exp\left(-\frac{(1+\alpha\epsilon)^2(1-\beta')^2(2\sqrt{2g})^2 \log^2 N}{2g \log N^{\beta'} + O(1)} \right) \to 0$$

since $\left|\Pi_{\beta'}\right| = O(N^{4(1-\beta')})$. So noticing that $\alpha(\gamma_m+\epsilon) = (1+\alpha\epsilon)$

$$\mathbf{P}(E) = o(1) + \mathbf{P}(E \cap D) \le o(1) + N^{-\rho(\alpha,\beta')-\delta} N^{4+4\beta'-4\alpha^2 F_{2,\beta'}(\gamma_m+\epsilon)+\delta/2}$$

if h is close to 1. $4+4\beta'-4\alpha^2 F_{2,\beta'}(\gamma_m+\epsilon) \xrightarrow{\epsilon\to 0} \rho(\alpha,\beta')$, thus $\mathbf{P}(E) \to 0$.

Proof of Theorem 1.14.

Lower bound. We recall the notation used in the proof of Theorem 1.10 by N. Kurt. For $\alpha \in (1/2, 1)$ we choose $1 \le k \le K+1$ such that

$$\alpha_k := \frac{\alpha(K-k+1)}{K} > \frac{1-\eta}{2} - \delta \tag{1.25}$$

(δ must be thought small). Let us now define recursively $\Gamma_{\alpha_1} := \Pi_{\alpha_1}$. Then for $i \geq 2$, we set Γ_{α_i} as follows: for any $B \in \Gamma_{\alpha_{i-1}}$ define $\Gamma_{B,\alpha_i} := \{B' \in \Pi_{\alpha_i} : B' \subseteq B/2\}$. Then

$$\Gamma_{\alpha_i} := \bigcup_{B \in \Gamma_{\alpha_{i-1}}} \Gamma_{B,\alpha_i}.$$

We re-use the notation $\underline{B}^{(k)}$ for a sequence of boxes $B_1 \supseteq B_2 \supseteq \cdots \supseteq B_k$, $B_i \in \Gamma_{\alpha_i}$ for all $1 \leq i \leq k$. Finally

$$D_k := \left\{ \underline{B}^{(k)} : \varphi_{B_i} \geq (\alpha - \alpha_i)\lambda 2\sqrt{2g}(1 - 1/K)\log N, \, \forall\, 1 \leq i \leq K \right\},$$

$$C_k := \{|D_k| \geq n_k\}.$$

We denote the biggest box of $\underline{B}^{(k)}$ with $B_{1,k}$. Let B be a box of side length $N^{\alpha_k}/2$ centered in $B_{1,k}$. Let $n_k := N^{\kappa + 4\alpha(k-1)\frac{(1-\lambda)^2}{K}}$, where κ is the constant appearing in [58, Lemma 3.2]. Define moreover for $\epsilon > 0$

$$A := \bigcup_{y \in B} \left\{ |\mathbf{E}(\varphi_y - \varphi_{x_B} | \mathscr{F}_{\alpha_k})| \geq 2\sqrt{2g}\epsilon(\alpha - \alpha_k)(1 - \gamma_K)\log N \right\}.$$

By Lemma 1.21 $\mathbf{P}(A^c) \to 1$ and $\mathbf{P}(C_k) \to 1$ as in Theorem 1.10 (C_k is the same event). So

$$\begin{aligned}
&\mathbf{P}\left(D_N(\eta) \leq N^{\frac{1-\eta}{2} - \delta} \right) \leq \\
&\leq o(1) + \\
&+\mathbf{P}\left(C_k \cap A^c \cap \left\{ \min_{y \in B} \varphi_y \geq 2\sqrt{2g}\eta \log N \right\} \right) \leq \\
&\stackrel{Def.of\, A, D_k}{\leq} o(1) + \\
&+\mathbf{P}\left(\min_{y \in B}(\varphi_y - \mathbf{E}(\varphi_y | \mathscr{F}_{\alpha_k})) \leq \right. \\
&\left. 2\sqrt{2g}\log N(\eta - (\alpha - \alpha_k)(1 - \gamma_K)(1 - \epsilon)) \right) \leq \\
&\leq \mathbf{P}\left(\max_{y \in V^{1/4}_{N^{\alpha_k}}} \varphi_y \geq 2\sqrt{2g}\log N(-\eta + (\alpha - \alpha_k)(1 - \gamma_K)(1 - \epsilon)) \right)
\end{aligned}$$

where in the latter inequality we used the fact that $V^{1/4}_{N^{\alpha_k}} \supseteq B$. For

$$2\sqrt{2g}\log N(-\eta + (\alpha - \alpha_k)(1 - \gamma_K)(1 - \epsilon)) > 2\sqrt{2g}\log N^{\alpha_k} \tag{1.26}$$

we would obtain thanks to Theorem 1.10 that for N large this probability tends to 0. But (1.25) and (1.26) give rise to a system of equations which has a solution for large K and N, α close to 1 and ϵ small when $1/2 + \eta/2k/K < \eta/2 + \delta + 1/2$.

Upper bound. We set $\theta := \frac{1-\eta}{2}$, $\beta := \theta + \delta$. We have first of all that

$$\mathbf{P}\left(\bigcup_{B\in\Pi_\beta}\{\varphi_B \geq 2\sqrt{2g}(1-\theta)\log N\}\right) \overset{N\to+\infty}{\longrightarrow} 0 \tag{1.27}$$

since we have the variance bounds and (1.15). Furthermore let us define

$$F := \left\{\bigcap_{B\in\Pi_\beta}\{\varphi_B \leq 2\sqrt{2g}(1-\theta)\log N\}\right\},$$
$$C := \left\{\bigcup_{B\in\Pi_\beta}\{\forall\, x \in B(\varphi_x \geq 2\sqrt{2g}\eta\log N)\}\right\}.$$

We then have

$$\begin{aligned}\mathbf{P}\left(D_N(\eta)\right. &\geq \left. N^{\theta+2\delta}\right) \leq \mathbf{P}(C) \leq \\ &\leq \mathbf{P}(F^c) + \mathbf{P}(F\cap C) \leq \\ &\overset{(1.27)}{\leq} o(1) + \mathbf{E}(\mathbf{P}(C|\mathscr{F}_\beta)\mathbf{1}_F).\end{aligned}$$

If $B \in \Pi_\beta$ we indicate with $B^{(1/4)}$ the sub-box $B(x_B, N^\beta/2)$. Choose $\epsilon > 0$ and define

$$A := \bigcup_{B\in\Pi_B}\bigcup_{y\in B^{(1/4)}}\left\{|\mathbf{E}(\varphi_y - \varphi_{x_B}|\mathscr{F}_{\partial_2 B})| \geq 2\sqrt{2g}\epsilon\log N\right\}.$$

With Lemma 1.21 we obtain that $\mathbf{P}(A)$ tends to 0 as in Theorem 1.11. We can further bound

$$\mathbf{P}(D_N(\eta) \geq N^{\theta+2\delta}) \leq o(1) + \mathbf{E}(\mathbf{P}(C|\mathscr{F}_\beta)\mathbf{1}_{F\cap A^c}).$$

To go on we notice that

$$\mathbf{P}(C|\mathscr{F}_\beta) \leq \left(\frac{N}{N^\beta}\right)^4 \max_{B\in\Pi_\beta}\mathbf{P}\left(\forall\, x \in B(\varphi_x \geq 2\sqrt{2g}\eta\log N)\right) \tag{1.28}$$

and in particular on $F\cap A^c$

$$\begin{aligned}&\mathbf{P}\left(\forall\, x\in B(\varphi_x \geq 2\sqrt{2g}\eta\log N)\right) \leq \\ &\leq \mathbf{P}\left(\forall\, x \in B(\varphi_x - \mathbf{E}(\varphi_x|\mathscr{F}_\beta) \geq 2\sqrt{2g}\log N(\eta-(1-\theta+\epsilon)))|\mathscr{F}_\beta\right) = \\ &= \mathbf{P}\left(\max_{x\in V^{1/4}_{N^\beta}}\varphi_x \leq 2\sqrt{2g}\log N(\theta+\epsilon)\right).\end{aligned}$$

By Theorem 1.8 this quantity is $O\left(\exp\left(-d\log^2 N\right)\right)$ for a positive d when for instance $\beta > (\theta + \epsilon)$ which implies $\epsilon < \delta$. To sum up

$$\mathbf{P}(C|\mathscr{F}_\beta) \leq \exp\left(2(1-\beta)\log N - d\log^2 N\right) \to 0$$

and recalling (1.28) we finish the proof.

Chapter 2.

The Gaussian Free Field

2.1. The d-dimensional Gaussian Free Field

The d-dimensional Gaussian free field (GFF) is a natural d-dimensional-time analog of Brownian motion. It has received great attention both in the physics community as a key tool to develop conformal field theories, especially in dimension 2 ([50]), and in mathematics, where it has shown to possess remarkable features allowing to connect different areas such as complex analysis and probability theory ([28, 52, 70, 72]). We will review here its construction and some of its main characteristics.

We begin by asking ourselves a question: is it possible to construct a standard Gaussian random variable on an infinite dimensional space? We already know that if H is isomorphic to $\mathbb{R}^k$ and $(\vec{e}_j)_{j=1}^k$ is a basis of the space, then any Gaussian $X \sim \mathcal{N}(0, \|x\|_H^2)$ can be written as

$$X = \sum_{j \le k} a_j \vec{e}_j,$$

where the a_j's are iid standard normal variables with law

$$\mathcal{W}(\mathrm{d}h) = \frac{1}{(2\pi)^{k/2}} \exp\left(- \|h\|_H^2 /2\right) \lambda(\mathrm{d}h)$$

absolutely continuous with respect to the Lebesgue measure λ. If we try to take this example over to an infinite dimensional (and say separable Hilbert) space H, possessing an orthonormal basis $(h_j)_{j\ge1}$, the formal sum

$$\sum_{j\ge1} a_j h_j \tag{2.1}$$

does not converge. The reason is well-known: if it did, then, for any orthonormal basis $(h_j)_{j\ge1}$, the random variables $h \in H \mapsto X_j(h) := (h, h_j)_H$ would be independent, standard normal random variables and therefore, by the strong law of large numbers, $\|h\|^2 = \sum_j X_j(h)^2$ would be infinite for $\mathcal{W}$-almost every h. One can circumvent this problem by allowing the series to converge in a larger space, call it Θ, which brings us to the definition of an abstract Wiener space. The construction of the abstract Wiener space was in fact first introduced by [43] to allow the construction of "infinite dimensional" Gaussian variables.

Definition 2.1 (Abstract Wiener space, [75])**.** *An* abstract Wiener space *is a triple* $(\Theta, H, \mathcal{W})$, *where*

- Θ *is a separable Banach space,*
- H *is a Hilbert space which is continuously embedded as a dense subspace of* Θ, *equipped with the scalar product* $(\cdot,\cdot)_H$,
- $\mathcal{W}$ *is a Gaussian probability measure on* Θ *satisfying*

$$\mathsf{E}_{\mathcal{W}}\left[\exp\left(i\left\langle \cdot, x^*\right\rangle\right)\right] = \exp\left(-\frac{\|h_{x^*}\|_H^2}{2}\right). \tag{2.2}$$

To understand (2.2) we first need to introduce the dual space Θ^*, the space of continuous linear functionals $x^* : \Theta \to \mathbb{R}$. If $\langle \cdot, x^* \rangle$ denotes the action of x^*, by Riesz's representation theorem for all x^* there exists a unique $h_{x^*} \in H$ for which for all $h \in H$ it holds that $\langle h, x^* \rangle = (h, h_{x^*})_H$. We can equip Θ with the σ-algebra $\mathcal{B}_\Theta$, which is the smallest σ-algebra with respect to which all the maps $\theta \mapsto \langle \theta, x^* \rangle$ are measurable. Then $\mathcal{W}$ is such that for all $x^* \in \Theta^*$ its Fourier transform fulfills

$$\widehat{\mathcal{W}}(x^*) = \exp\left(-\frac{\|h_{x^*}\|_H^2}{2}\right). \tag{2.3}$$

Wiener's construction of Brownian motion can be viewed as the original case in which H is the Hilbert space of absolutely continuous $h : [0,1] \to \mathbb{R}$ such that $h(0) = 0$ and $\dot{h} \in L^2([0,1], \mathbb{R})$ with the norm $\|h\| = \|\dot{h}\|_{L^2([0,1],\mathbb{R})}$. Then $\Theta = L^2([0,1], \mathbb{R})$.
At this point we have to introduce another fundamental object for our work, the *Paley-Wiener integral* $\mathcal{I}(h)$. $\mathcal{I}$ is viewed as a mapping

$$\begin{aligned} \mathcal{I} :\ x^* \in \Theta^* \quad &\mapsto \quad \mathcal{I}(h_{x^*}) \in L^2(\mathcal{W}) \\ &\qquad \theta \in \Theta \mapsto [I(h_{x^*})](\theta) := \langle \theta, x^* \rangle . \end{aligned}$$

Therefore $\mathcal{I}$ is an isometry from $\{h_{x^*} : x^* \in \Theta^*\} \to L^2(\mathcal{W})$, and since the former set is dense in H, it admits a unique extension to the whole of H.

Example 2.2 (2-dimensional GFF, [40, 73])**.** *We would like to define the 2-dimensional GFF on* $\mathbb{D} := [0,1]^2 \subseteq \mathbb{R}^2$ *with Dirichlet boundary conditions in the above framework. Denote as* $H_s(\mathbb{D})$ *the set of smooth, real-valued functions on* $\mathbb{R}^2$ *that are supported on a compact subset of* $\mathbb{D}$. *Equip this space with the Dirichlet inner product*

$$(f,g)_\nabla := \int_{\mathbb{D}} \nabla f(x) \cdot \nabla g(x) \mathrm{d}x.$$

We set $H := H(\mathbb{D})$ to be the closure of $H_s(\mathbb{D})$ with respect to this inner product (this yields the so-called Sobolev space $H_0^1(\mathbb{D})$). The GFF h is defined as the family

$$\{(h, f)_\nabla : f \in H(\mathbb{D})\}.$$

It is a Gaussian family with covariances

$$\mathrm{Cov}\left((h, f)_\nabla, (h, g)_\nabla\right) = (f, g)_\nabla.$$

It is possible to write down explicitly an orthonormal basis for $H(\mathbb{D})$, which is indexed by $j, k \in \mathbb{N} \setminus \{0\}$,

$$\vec{e}_{j,k} := \frac{2\sqrt{2}\sin(j\pi x)\sin(k\pi x)}{\sqrt{j^2 + k^2}}.$$

Hence the GFF h displays the alternative representation

$$h = \sum_{j,k \in \mathbb{N} \setminus \{0\}} a_{j,k} \vec{e}_{j,k}$$

where $a_{j,k}$ are i. i. d. standard Gaussians and the series is convergent in $H^{-1}(\mathbb{D})$ ([28]).

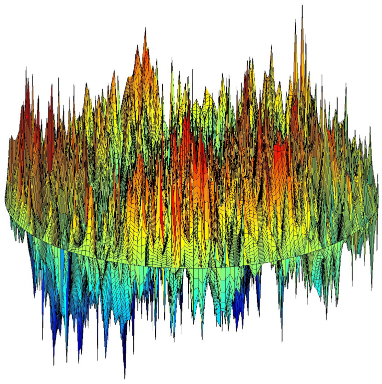

Figure 2.1.: 2-dimensional GFF on the unitary disk with Dirichlet boundary conditions. Picture by Nam-Gyu Kang.

It is easy now to see that in an abstract Wiener space $(\Theta, H, \mathcal{W})$ the series (2.1) is almost surely finite. Indeed, using separability, one can show that $\mathcal{B}_\Theta$ is contained in the $\mathcal{W}$-completion of $\sigma\left(\bigcup_{n \geq 1} \mathscr{F}_n\right)$, $\mathscr{F}_n$ being the sigma-algebra generated by $\left\{I[(h_j)], j \leq n\right\}$. Being a Gaussian family this allows to say that

$$\theta \mapsto S_n := \sum_{j=1}^{n} I[(h_j)](\theta) h_j$$

is the $\mathcal{W}$-conditional expectation of θ given $\mathscr{F}_n$. Doob's martingale convergence theorem allows to conclude that $\theta = \lim_{n\to+\infty} S_n$. This Wiener series representation yields a number of interesting consequences, among which the *Cameron-Martin formula*.

Theorem 2.3 ([75]). *Let $(\Theta, H, \mathcal{W})$ be an abstract Wiener space. If $h \in H$ and*

$$R_h(\theta) = \exp\left([I(h)](\theta) - \frac{\|h\|_H^2}{2}\right)$$

then the distribution of $\theta \mapsto \theta + h$ under $\mathcal{W}$ is absolutely continuous with respect to $\mathcal{W}$ and R_h is the corresponding Radon–Nikodym derivative.

From the standpoint of Wiener series, this observation comes down to the fact that if one translates the standard Gauss distribution γ on $\mathbb{R}$ by $a \in \mathbb{R}$, then the translated measure is absolutely continuous with respect to γ and has Radon–Nikodym derivative $\exp\left(ax - \frac{a^2}{2}\right)$.

2.2. Multiplicative chaos

Multiplicative chaos represents a more general framework in which one can view the construction of the GFF. The definition stands as follows:

Definition 2.4. *In dimension d a* standard Gaussian multiplicative chaos *is a random measure on a given domain $D \subseteq \mathbb{R}^d$ that can be written for any $A \in \mathcal{B}(D)$ as*

$$M_\gamma(A) = \int_A \exp\left(\gamma X(x) - \frac{\gamma^2}{2}\mathbb{E}\left(X^2(x)\right)\right)\sigma(\mathrm{d}x)$$

with X a centered Gaussian field and σ a Radon measure on D.

The case where X is the Gaussian Free Field (either in 2 or 4 dimensions) represents an instance of multiplicative chaos. As we have seen the definition is ill-posed because X turns out to be a random distribution. The first description of this object to our knowledge goes back to Kahane [49] who considered in his original paper a field X whose covariances were represented by the positive kernel

$$K(x, y) := \mathbb{E}\left(X(x)X(y)\right) = \ln_+ \frac{1}{|x-y|} + g(x, y) \tag{2.4}$$

with $\ln_+$ the positive part of ln and g a continuous bounded function on D^2. It appears that different instances of multiplicative chaos share the same universal features, and we refer to the nice survey [69] for a complete description. In particular, we point out here that he considered the analog of the (nowadays called) thick points of the GFF that we analyze in our work, and obtained the counterparts (see for a more precise statement [69, Theorems 4.1 and 4.2]) of the results of [46] for the two-dimensional GFF and of this chapert.

2.2.1. Behavior of the maximum

There are more features that the class of log-correlated fields seems to possess, and among these there is the behavior of the maximum. Recently [62] showed the convergence of the maximum of the log-correlated field of Kahane to a Gumbel distribution convoluted by the limit of the derivative martingale. The analog limit in the discrete setting was recovered by [15] for the DGFF and the corresponding statistics of local maxima were shown by [12].

2.3. Thick points for a Gaussian Free Field in 4 dimensions

This section is concerned with the study of fractal properties of thick points for 4-dimensional Gaussian Free Field. We adopt the definition of Gaussian Free Field on $\mathbb{R}^4$ introduced by [20] viewed as an abstract Wiener space with underlying Hilbert space $H^2(\mathbb{R}^4)$. We can prove that for $0 \leq a \leq 4$, the Hausdorff dimension of the set of a-high points is $4-a$. We also show that the thick points give full mass to the Liouville Quantum Gravity measure on $\mathbb{R}^4$. This section is based on a joint paper with Rajat Subhra Hazra [22] which has been submitted.

2.3.1. Introduction

The main goal in constructing *Liouville quantum gravity* (LQG) measure is to construct a random metric on a Riemannian manifold (equipped with an Euclidean metric $\mathrm{d}\omega$), which has a form $e^{\gamma X(\omega)}\mathrm{d}\omega$ where $X(\omega)$ is a Gaussian free field and γ is a coupling parameter. The KPZ relation formulated by Knizhnik, Polyakov and Zamolodchikov [54] gives the relation between volume exponents derived using the abovementioned quantum metric and the Euclidean metric respectively. They may also be considered in terms of measures and under this light lots of progress has been made recently. In dimension 2, in the breakthrough work of [29] the existence of the Liouville quantum gravity measure was proved and the KPZ relation was derived. For an overview on ongoing activity see the reviews [40], [69]. The KPZ relation holds also in higher dimensions. In a recent article by [20] an analogue of the LQG measure was introduced in dimension 4. The aim of this article is to study some fractal properties of the Gaussian free field (GFF) in $\mathbb{R}^4$ and relate it to the LQG.

We focus our attention on the so-called *thick points*, which roughly speaking are points unusually high. In an interesting work, [46] showed that the Hausdorff dimension of the set of a-thick points is $2 - a$ for $0 \le a \le 2$ for the planar GFF. The set of thick points is important in understanding the support of two dimensional Liouville quantum gravity. It was shown in fact in [29] that the LQG measure is almost surely supported on the thick points. In general though the definition is ill-posed, so a regularization (the circle average process) is needed to define it. Motivated by the definition of sphere average introduced by [20] in dimensions 4, we study the set of thick points for the sphere average process of the four-dimensional Gaussian Free field.

The Euclidean Gaussian Free Field on $\mathbb{R}^4$ is denoted by means of the abstract Wiener space (see Definition 2.1), where the underlying Hilbert space is $H^2(\mathbb{R}^4)$ equipped with the inner product given by $\langle (I - \Delta)^2 \cdot, \cdot \rangle$. Note that, unlike the standard definition of GFF (see [73]) we work with a Bilaplacian (or Bessel operator) and the second order Sobolev spaces. Other than Chen and Jakobson's recent article the main motivation for considering this model comes from the discrete analogue of this GFF which turns out to be related to the membrane model (cf. [58]) defined on $\mathbb{Z}^d$. It is known that in dimension 4 the model undergoes a phase transition in terms of the behavior of the Gibbs volume measure, as was proved in [57]. So dimension 4 turns out to be intrinsically interesting. Recently, some work on the fractal dimension of the thick points in this discrete setting has been carried through by [23] for the 2-D discrete Gaussian Free Field and [21] on the discrete 4-D membrane model. We note however that since the GFF we are dealing with is defined on the whole space, an appropriate definition of the GFF on bounded domains still remains an open problem. Thick points are an object interesting in itself, as they reveal the structure of the exceedances of the GFF. Even more interestingly though, they represent, as we will see, the support of the LQG: this random measure hence presents a support with large flat (zero-mass) areas and some sparse large spikes. It was pointed out in the recent survey by [69] (Theorem 4.1) that in Kahane's seminal work ([48]) on Gaussian multiplicative chaos similar results on the Hausdorff dimension of thick points are proved. It was also pointed out that the circle average process has a similar multifractal behavior and our result provides another confirmation for this.

In this article we show the manner in which the thick points are related to the LQG measure in dimension 4. We prove in Theorem 2.6 that the set of thick points gives full mass to the LQG measure. We then study the multifractal properties of the set of thick points. In particular, we show in Theorem 2.8 that the set of a-thick points has Hausdorff dimension $4 - a$ when $0 \le a \le 4$. When $a > 4$, the set of thick points is almost surely empty. We point out that results of similar nature were known for Gaussian multiplicative chaos (in any dimension). For generalizations and further references of such results we refer to [69]. More precisely the outline of the article is as follows. In Section 2.3.2 we recall the model introduced by [20] and state our main result more precisely. In Section 2.3.3 we list some basic properties of the sphere average process and also provide a proof of Theorem 2.6 using a so-called *rooted measure*. The proof of Theorem 2.8 is given in Sections 2.3.6 and 2.3.7 and relies on proving two different bounds. For the upper bound we use the version of the Kolmogorov-Centsov theorem derived by [46]. For the lower bound we use a standard finite-energy method and the Markov property of the GFF.

2.3.2. GFF model and statement of the main results

To keep the paper self contained we review in this section some definitions of the GFF on $\mathbb{R}^4$ from [20] and state some properties of the sphere average process which will be useful in deriving our main result. In first place, in defining the abstract Wiener space for the GFF, although the introduction of the set Θ is evidently important for its definition, its choice is not unique as explained in [75], Corollary 8.3.2 and afterwards. Moreover $\mathcal{W}(H) = 0$, as it follows from the discussion in Section 2.1 that $\mathcal{W}$-a. s. $\|h\|_H^2 = +\infty$ for $h \in \Theta$ and clearly $\|h\|_H^2 < +\infty$ for $h \in H$. In our setting, we consider the underlying Hilbert space to be $H := H^2\left(\mathbb{R}^4\right)$ which is the completion of the Schwartz space $\mathcal{S}\left(\mathbb{R}^4\right)$ equipped with the inner product

$$(f_1, f_2)_H = \int_{\mathbb{R}^4} (I - \Delta)^2 f_1(x) f_2(x)\,\mathrm{d}x \quad \text{for all } f_1, f_2 \in \mathcal{S}\left(\mathbb{R}^4\right).$$

$H^{-2}\left(\mathbb{R}^4\right)$ is the Hilbert space consisting of tempered distributions μ such that

$$\|\mu\|_{H^{-2}}^2 = \frac{1}{(2\pi)^4} \int_{\mathbb{R}^4} \left(1 + |\xi|^2\right)^{-2} |\hat{\mu}(\xi)|^2 \,\mathrm{d}\xi < \infty.$$

where $\hat{\mu}$ is the Fourier transform. It is possible to identify H with H^{-2} through the linear isometry $(I - \Delta)^{-2} : H^{-2} \to H$. By abuse of notation we will call h_ν the image of $\nu \in H^{-2}$ under $(I - \Delta)^{-2}$, that is, h_ν is the unique element in H such that $\langle h, \nu\rangle = (h, h_\nu)_H$ for all $h \in H$. By (2.2), we have $\left\{\mathcal{I}(h_\nu) : \nu \in H^{-2}\right\}$ is also a Gaussian family whose covariance is given by

$$\mathsf{E}_{\mathcal{W}}\left[\mathcal{I}(h_{\nu_1})\, \mathcal{I}(h_{\nu_2})\right] = \langle h_{\nu_1}, h_{\nu_2}\rangle_H = \langle \nu_1, \nu_2\rangle_{H^{-2}}.$$

For every $x \in \mathbb{R}^4$ and $\epsilon > 0$ denote as $\sigma_\epsilon^x \in H^{-2}$ the tempered distribution given by

$$\langle f, \sigma_\epsilon^x\rangle = \frac{1}{2\pi^2\epsilon^3} \int_{D(x,\epsilon)} f(y)\,\mathrm{d}\sigma(y), \quad \text{for all } f \in \mathcal{S}\left(\mathbb{R}^4\right),$$

where $\mathrm{d}\sigma$ is the surface area measure on $D(x, \epsilon)$, the sphere of radius ϵ around x. Interestingly, [20] noted that $\left\{\mathcal{I}\left(h_{\sigma_\epsilon^x}\right) : \epsilon > 0\right\}$ fails to possess the Markov property and considered the following Gaussian family:

$$\left\{\mathcal{I}\left(h_{\sigma_\epsilon^x}\right), \mathcal{I}\left(h_{\mathrm{d}\sigma_\epsilon^x}\right) : x \in \mathbb{R}^4, \epsilon > 0\right\},$$

where $\mathrm{d}\sigma_\epsilon^x$ the tempered distribution given by $\langle f, \mathrm{d}\sigma_\epsilon^x\rangle := \frac{\mathrm{d}}{\mathrm{d}\epsilon}\langle f, \sigma_\epsilon^x\rangle$ for all $f \in \mathcal{S}\left(\mathbb{R}^4\right)$. It is important to point out at this juncture that such a collection is reminiscent of the double boundary conditions needed for the membrane model in the discrete case ([56]). Let $\zeta := (1,1)^T$ and

$$\mathbf{B}(r) := \begin{pmatrix} I_1(r)/r & I_1'(r) \\ I_2(r)/r & I_1''(r) \end{pmatrix},$$

where I_k are the modified Bessel functions of order $k \in \mathbb{N}$. Define

$$\mu_\epsilon^x := \zeta^\top \mathbf{B}^{-1}(\epsilon) \begin{pmatrix} \sigma_\epsilon^x \\ \mathrm{d}\sigma_\epsilon^x \end{pmatrix}. \tag{2.5}$$

[20] showed that $\mu_\epsilon^x \in H^{-2}(\mathbb{R}^4)$ and $\left\{\mathcal{I}\left(h_{\mu_\epsilon^x}\right) : x \in \mathbb{R}^4, \epsilon > 0\right\}$ forms a Gaussian family with the correct Markovian properties and is the suitable candidate for the sphere average process.

Definition 2.5 (Thick points of the sphere average). *For the sphere average process the set of a-thick points is defined as*

$$T(a) = \left\{x \in \mathbb{R}^4 : \lim_{\epsilon \to 0} \frac{\mathcal{I}\left(h_{\mu_\epsilon^x}\right)}{\sqrt{2\pi^2 G(\epsilon)}} = \sqrt{2a}\right\}. \tag{2.6}$$

Here $G(\epsilon) = \mathsf{Var}_{\mathcal{W}}\left(\mathcal{I}\left(h_{\mu_\epsilon^x}\right)\right)$ *and an explicit expression using Bessel functions is given in* (2.9)*.*

We would also need a definition of another set quite similar to the above:

$$T_{\geq}(a) = \left\{x \in \mathbb{R}^4 : \limsup_{\epsilon \to 0} \frac{\mathcal{I}\left(h_{\mu_\epsilon^x}\right)}{\sqrt{2\pi^2 G(\epsilon)}} \geq \sqrt{2a}\right\}. \tag{2.7}$$

It is easy to see that

$$T(a) \subset T_{\geq}(a).$$

One of the important results of [20] (Theorem 5) was to show the existence of the Liouville quantum gravity measure in $\mathbb{R}^4$. Define a random measure on $\mathbb{R}^4$ by

$$m_\epsilon^\theta(\mathrm{d}x) = E_\epsilon^\theta(x)\mathrm{d}x,$$

where
$$E_{\epsilon}^{\theta} = \exp\left(\gamma \mathcal{I}\left(h_{\mu_{\epsilon}^{x}}\right) - \frac{\gamma^2}{2} G(\epsilon)\right).$$
If $\epsilon_n = \epsilon_0^n$ with $\epsilon_0 \in (0,1)$ and $0 < \gamma^2 < 2\pi^2$, then there exists a non-negative measure m^θ on $\mathbb{R}^4$ such that the following convergence holds for every $f \in C_c(\mathbb{R}^4)$:
$$\int_{\mathbb{R}^4} f(x) m_{\epsilon_n}^{\theta}(\mathrm{d}x) \to \int_{\mathbb{R}^4} f(x) m^{\theta}(\mathrm{d}x) \text{ as } n \to \infty \tag{2.8}$$
$\mathcal{W}$-almost surely and also in $L^2(\mathcal{W})$. It is also known that this measure is almost surely positive.

In the following Theorem we show that the set of thick points gives full measure to the LQG measure in $\mathbb{R}^4$.

Theorem 2.6. *Let $0 < \gamma^2 < 2\pi^2$, then for $a = \gamma^2/4\pi^2$ we have*
$$m^\theta(T(a)^c) = 0 \ \ \mathcal{W} - a.s.$$
That is, the set $T(a)$ gives full mass to the measure $m^\theta(\cdot)$.

For the proof of Theorem 2.6 we construct the *rooted measure or Peyrière measure*. For the use of rooted measures see [29, 69].

Before we state our main result on fractal properties of thick points, we recall the definition of Hausdorff dimension and Hausdorff measure.

Definition 2.7 (Hausdorff dimension)**.** *Let X be a metric space and $S \subseteq X$. For every $d \geq 0$ and $\delta > 0$ define the Hausdorff-d-measure in the following way:*
$$C_\delta^d(S) := \inf\Big\{\sum_i \mathrm{diam}(E_i)^d \ : \ E_1, E_2, E_3, \ldots, \textit{cover } S, \ \mathrm{diam}(E_i) \leq \delta\Big\},$$
i.e. we are considering coverings of S by sets of diameter no more than δ. Then
$$C_{\mathcal{H}}^d(S) = \sup_{\delta > 0} C_\delta^d(S) = \lim_{\delta \downarrow 0} C_\delta^d(S)$$
is the Hausdorff-d-measure of the set S. The Hausdorff dimension *of X is defined by*
$$\dim_{\mathcal{H}}(X) := \inf\{d \geq 0 : \ C_{\mathcal{H}}^d(X) = 0\}.$$
For S subset of a metric space X we would denote by $\dim_{\mathcal{H}}(S)$ as the Hausdorff dimension of S and $C_{\mathcal{H}}^{\alpha}(S)$ will denote the Hausdorff-α-measure of the set S.

Theorem 2.8. *For $0 \leq a \leq 4$, the Hausdorff dimension of $T(a)$ is $4 - a$. For $a > 4$, we have that $T(a)$ is empty.*

Remark 2.9. *The above result shows similarity with the membrane model. In [21] it was shown that discrete fractal dimension of the a-high points is $4 - 4a^2$.*

To prove Theorem 2.8 we apply some of the techniques implemented in [26, 25] to show similar results for occupation measures of planar or spatial Brownian motion.

2.3.3. GFF model and some estimates

This section is devoted to providing some details about the behavior of the sphere average process, such as the covariance structure. We then use them to derive a proof of Theorem 2.6.

2.3.4. Some more properties of the sphere average process: covariance structure

Let us denote as $D(0, R)$ the hypersphere centered at 0 with radius $R > 0$. Let I_r, K_r be the modified Bessel functions of order $r \in \mathbb{N} \cup \{0\}$. Define the positive function $G : (0, \infty) \mapsto (0, \infty)$ by

$$G(r) \quad := \left(-\tfrac{1}{4\pi^2}\right) \tfrac{2I_1(r)K_1(r)+2I_2(r)K_0(r)-1}{I_1^2(r)-I_0(r)I_2(r)} \ . \tag{2.9}$$

It can be shown that G is strictly decreasing and smooth, with $\lim_{r\to 0} G(r) = +\infty$ and $\lim_{r\to+\infty} G(r) = 0$. It also follows from the properties of the Bessel functions that as r decreases to 0, $G(r)$ asymptotically behaves like $-\frac{1}{2\pi^2}\log r$. Then, we have that

(a) given $x \in \mathbb{R}^4$ and $\epsilon_1 \geq \epsilon_2 > 0$,

$$\mathsf{E}_{\mathcal{W}}\left[\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right)\mathcal{I}\left(h_{\mu^x_{\epsilon_2}}\right)\right] = \mathsf{E}_{\mathcal{W}}\left[\mathcal{I}^2\left(h_{\mu^x_{\epsilon_1}}\right)\right] = G(\epsilon_1). \tag{2.10}$$

(b) Given $x, y \in \mathbb{R}^4$, $x \neq y$, and $\epsilon_1, \epsilon_2 > 0$ with $\overline{D(x, \epsilon_1)} \cap \overline{D(y, \epsilon_2)} = \emptyset$,

$$\mathsf{E}_{\mathcal{W}}\left[\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right)\mathcal{I}\left(h_{\mu^y_{\epsilon_2}}\right)\right] = \frac{1}{2\pi^2}K_0(|x-y|), \tag{2.11}$$

where K_0 is the modified Bessel function of order 0.

(c) Given $x, y \in \mathbb{R}^4$, $x \neq y$, and $\epsilon_1, \epsilon_2 > 0$ with $D(y, \epsilon_2) \subseteq D(x, \epsilon_1)$,

$$\mathsf{E}_{\mathcal{W}}\left[\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right)\mathcal{I}\left(h_{\mu^y_{\epsilon_2}}\right)\right] = I_0(|x-y|)\,G(\epsilon_1) - \frac{1}{4\pi^2}\frac{I_2(|x-y|)}{I_1^2(\epsilon_1) - I_0(\epsilon_1)\,I_2(\epsilon_1)}. \tag{2.12}$$

The next lemma states one of the most useful and important properties of the spherical average process and is analogous to the properties of the two dimensional circular average process studied in [29, 46]. It shows that for fixed $x \in \mathbb{R}^4$, the spherical average after a time change is a Brownian motion and in disjoint annuli two such motions are independent. We briefly sketch the proof of the following lemma as it is an easy consequence after one compares the covariance structure.

Lemma 2.10. *(a) Let $G(\cdot)$ be as in (2.9) and for $x \in \mathbb{R}^4$, let $B(x,t) = \mathcal{I}\left(h_{\mu^x_{G^{-1}(t)}}\right)$. Then $B(x,t) - B(x,t_1)$ has the same distribution as a standard Brownian motion for $t \geq t_1$.*

(b) Given $x, y \in \mathbb{R}^4$ and $t_1 \leq t \leq t_2$ and $s_1 \leq s \leq s_2$ be such that $D(x,s_2) \setminus D(x,s_1)$ and $D(y,t_2) \setminus D(y,t_1)$ are disjoint, then $\{B(x,s) - B(x,s_1)\}_{s_1 \leq s \leq s_2}$ is independent of $\{B(y,t) - B(y,t_1)\}_{t_1 \leq t \leq t_2}$.

Proof. (a) It follows from (2.10) that for $t_1 \leq s \leq t$ one has

$$\begin{aligned}&\mathrm{cov}_{\mathcal{W}}(B(x,t) - B(x,t_1), B(x,s) - B(x,t_1)) = \\ &= G(G^{-1}(s)) - G(G^{-1}(t_1)) - G(G^{-1}(t_1)) + G(G^{-1}(t_1)) = s - t_1\end{aligned}$$

Here we have used the fact that $G(\cdot)$ and $G^{-1}(\cdot)$ are decreasing functions and hence, as $t_1 \leq s \leq t$ we have $G^{-1}(t_1) \geq G^{-1}(s) \geq G^{-1}(t)$.

(b) As the annuli are disjoint it follows that $|x - y| > t_2 + s_2 > t + s > t_1 + s_1$ and hence again using (2.11) we obtain

$$\mathrm{cov}_{\mathcal{W}}\left(B(y,t) - B(y,t_1), B(x,s) - B(x,s_1)\right) = 0.$$

□

2.3.5. Proof of Theorem 2.6

Let Γ be a compact subset of $\mathbb{R}^4$. Let $\mathcal{B}(\Gamma)$ be the Borel sigma algebra of subsets of Γ. We define a rooted measure on $\mathcal{B}(\Theta) \otimes \mathcal{B}(\Gamma)$ as

$$\mathcal{M}(\mathrm{d}x\mathrm{d}\theta) = \frac{m^\theta(\mathrm{d}x)\mathcal{W}(\mathrm{d}\theta)}{|\Gamma|}.$$

Here $|\Gamma|$ denotes the volume of the set Γ with respect to the Lebesgue measure. Note that $\mathcal{M}(\Theta \times \Gamma) = \mathsf{E}_{\mathcal{W}}\left[m^\theta(\Gamma)\right] |\Gamma|^{-1} = 1$ and as such $\mathcal{M}$ is a probability measure on the space $\Gamma \times \Theta$.

Let $r(t) := G^{-1}(t + G(R))$, $R > 0$ fixed and define

$$\widetilde{B}(x,t)(\theta) := \mathcal{I}\left(h_{\mu^x_{r(t)}}\right)(\theta) - \mathcal{I}\left(h_{\mu^x_R}\right)(\theta).$$

The following lemma allows us to view the random measure m^θ in a different way. We show that the joint distribution of $(x, \widetilde{B}(x,t))$ under $\mathcal{M}(\mathrm{d}x\mathrm{d}\theta)$ is nothing but the distribution of $(x, \widetilde{B}(x,t))$ under $\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x$ and in the latter case the marginal on Θ does not depend on x.

Lemma 2.11. *Let $0 < \gamma^2 < 2\pi^2$. For any compact set Γ and any $F \in C_c(\mathbb{R}^4 \times \mathbb{R})$ we have*

$$\int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t)(\theta))\mathcal{M}(\mathrm{d}x\mathrm{d}\theta) = \frac{1}{|\Gamma|}\int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t)(\theta) + \gamma t)\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x. \tag{2.13}$$

Proof. Note that for almost every θ, the map $x \in \Gamma \mapsto F(x, \widetilde{B}(x,t)(\theta))$ is continuous by Corollary 3 of [20]. So from the weak convergence in (2.8) we have that

$$\lim_{n\to\infty} \int_\Gamma F(x, \widetilde{B}(x,t))m^\theta_{\epsilon_n}(\mathrm{d}x) = \int_\Gamma F(x, \widetilde{B}(x,t))m^\theta(\mathrm{d}x).$$

Since the function in the integral is bounded we have for some constant C and for all n

$$\int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))m^\theta_{\epsilon_n}(\mathrm{d}x)\mathcal{W}(\mathrm{d}\theta) \leq C|\Gamma|.$$

So by dominated convergence

$$\lim_{n\to\infty} \frac{1}{|\Gamma|}\int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))m^\theta_{\epsilon_n}(\mathrm{d}x)\mathcal{W}(\mathrm{d}\theta) = \int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))\mathcal{M}(\mathrm{d}x\mathrm{d}\theta). \tag{2.14}$$

Note that for small enough $\epsilon > 0$

$$\mathrm{Cov}(\widetilde{B}(x,t), h_{\mu^x_\epsilon}) = G(r(t)) - G(R) = t$$

holds, so for n large enough we have by Cameron-Martin theorem

$$\begin{aligned}
\int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))m^\theta_{\epsilon_n}(\mathrm{d}x)\mathcal{W}(\mathrm{d}\theta) &= \int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))E^\theta_{\epsilon_n}(x)\mathrm{d}x\mathcal{W}(\mathrm{d}\theta) \\
&= \int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t)(\theta))E^\theta_{\epsilon_n}(x)\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x \\
&= \int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t)(\theta + \gamma h_{\mu^x_{\epsilon_n}})\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x \\
&= \int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t)(\theta) + \gamma t)\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x.
\end{aligned}$$

We have the required statement in the lemma using (2.14).

□

Proof of Theorem 2.6. Using the fact $\mathrm{E}_\mathcal{W}\left[m^\theta(A)\right] = |A|$ for any bounded set A it follows that the marginal of $\mathcal{M}$ on Γ is nothing but the normalized Lebesgue measure on Γ. Hence by Theorem 9.2.2. of [75] there exists a Borel measurable map

$$x \in \Gamma \to \mathcal{L}_x(\cdot) \in M_1(\Theta),$$

where $M_1(\Theta)$ is the set of probability measures on Θ and the following holds

$$\mathcal{M}(\mathrm{d}x\mathrm{d}\theta) = \mathcal{L}_x(\mathrm{d}\theta)\frac{\mathrm{d}x}{|\Gamma|}.$$

Note that $\mathcal{L}_x(\mathrm{d}\theta)$ is nothing but the regular conditional probability. Now using the above decomposition we have that

$$\int_\Theta \int_\Gamma F(x, \widetilde{B}(x,t))\mathcal{M}(\mathrm{d}x\mathrm{d}\theta) = \frac{1}{|\Gamma|}\int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t))\mathcal{L}_x(\mathrm{d}\theta)\mathrm{d}x.$$

So from (2.11) we have for any compact set Γ and $F \in C_c(\mathbb{R}^4 \times \mathbb{R})$

$$\frac{1}{|\Gamma|}\int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t))\mathcal{L}_x(\mathrm{d}\theta)\mathrm{d}x = \frac{1}{|\Gamma|}\int_\Gamma \int_\Theta F(x, \widetilde{B}(x,t) + \gamma t)\mathcal{W}(\mathrm{d}\theta)\mathrm{d}x. \tag{2.15}$$

If we denote μ_x to be law of $\widetilde{B}(x,t)$ under $\mathcal{L}_x(\mathrm{d}\theta)$ and ν be the law $\widetilde{B}(x,t) + \gamma t$ under $\mathcal{W}(\mathrm{d}\theta)$ on $\mathbb{R}$ it is possible to see that ν is the law of a standard Brownian motion with a drift. Since (2.15) holds for any compact set Γ, it is easy to show that for almost every $x \in \mathbb{R}^4$, $\mu_x = \nu$. If we take $a = \gamma^2/4\pi^2$ and use the fact that the sphere average process is a time inversion of a Brownian motion (see Lemma 2.10), then the set of thick points can also be written as

$$T(a) = \left\{ x \in \mathbb{R}^4 : \lim_{t\to\infty} \frac{\widetilde{B}(x,t)}{t} = \gamma \right\}.$$

Now from the discussion above we have that

$$\mathcal{M}(T(a)^c) = \frac{1}{|\Gamma|}\int_\Gamma \mathcal{L}_x(T(a)^c)\mathrm{d}x.$$

Since the law of $\widetilde{B}(x,t)$ under $\mathcal{L}_x$ is the same as the law of Brownian motion with a drift, the condition for the thick points gets satisfied with probability 1. So we have $\mathcal{M}(T(a)^c) = 0$, which together with the fact that $m^\theta(\cdot)$ is a positive measure with probability 1 proves the result. □

2.3.6. Upper bound of Theorem 2.8

In this section we prove the upper bound. By the countable stability property, viz.

$$\dim_{\mathcal{H}}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sup_{1\le i\le\infty} \dim_{\mathcal{H}}(E_i)$$

it is enough to show that a.s. for $R \geq 1$,

$$\dim_{\mathcal{H}} T_{\geq}(a,R) = \dim_{\mathcal{H}} \left\{ x \in D(0,R) : \limsup_{\epsilon \to 0} \frac{\mathcal{I}\left(h_{\mu_\epsilon^x}\right)}{\sqrt{2\pi^2 G(\epsilon)}} \geq \sqrt{2a} \right\} \leq 4 - a. \quad (2.16)$$

Hence if we cover $\mathbb{R}^4$ with a countable union of balls of radius $R = 1, 2, \ldots$, this will prove the upper bound. The next proposition gives the local Hölder continuity of the process and through it we can determine a modification of the process which has some uniform estimates on the increments. It is similar to Proposition 2.1 of [46] and uses Lemma C.1 of [46]. The proof also uses some finer estimates on the covariance functions and some bounds on Bessel functions which are provided in Sec*A.2.

Proposition 2.12. *There exists a modification* $\widetilde{X}$ *of the process* $\{\mathcal{I}\left(h_{\mu_t^z}\right) : z \in D(0,R),\, t \in (0,1)\}$ *such that for every* $0 < \gamma < \frac{1}{2}$ *and* $\epsilon, \zeta > 0$ *there exists* $M > 0$ *such that the following holds:*

$$|\widetilde{X}(z,r) - \widetilde{X}(w,s)| \leq M \left(\log \frac{1}{r}\right)^{\zeta} \frac{|(z,r)-(w,s)|^{\gamma}}{r^{(1+\epsilon)\gamma}}, \quad (2.17)$$

for all $z, w \in D(0,R)$ *and* $r, s \in (0,1]$ *with* $1/2 \leq r/s \leq 2$.

Proof. Consider now x, $y \in D(0,R)$, ϵ_1, $\epsilon_2 \in (0,1)$ and we abbreviate

$$H_{\epsilon_1,\epsilon_2}(x,y) := \mathsf{Cov}_{\mathcal{W}}\left(\mathcal{I}\left(h_{\mu_{\epsilon_1}^x}\right), \mathcal{I}\left(h_{\mu_{\epsilon_2}^y}\right)\right).$$

We distinguish between three cases:

Case 1 Let $x = y$. By Lemma A.1, we have

$$\begin{aligned} |H_{\epsilon_1,\epsilon_1}(x,x) - H_{\epsilon_2,\epsilon_1}(x,x)| &\leq |H_{\epsilon_1,\epsilon_1}(x,x) - H_{\epsilon_1,\epsilon_2}(x,x)| \\ &\quad + |H_{\epsilon_2,\epsilon_1}(x,x) - H_{\epsilon_1,\epsilon_2}(x,x)| \\ &\overset{(2.10)}{\leq} |G(\epsilon_1) - G(\epsilon_1 \vee \epsilon_2)| + |G(\epsilon_2) - G(\epsilon_1 \vee \epsilon_2)| \\ &\leq C \frac{|\epsilon_1 - \epsilon_2|}{\epsilon_1 \wedge \epsilon_2}. \end{aligned}$$

Here we have used that $|\log(x/y)| \leq \frac{|x-y|}{x \wedge y}$.

Case 2 Let $\overline{D(x, \epsilon_1)} \cap \overline{D(y, \epsilon_2)} = \emptyset$. In this case $|x-y| > \epsilon_1 + \epsilon_2 > \epsilon_1$. Then

$$\begin{aligned} |H_{\epsilon_1,\epsilon_1}(x,x) - H_{\epsilon_1,\epsilon_2}(x,y)| &= \left| G(\epsilon_1) - \frac{1}{2\pi^2} K_0(|x-y|) \right| \\ &\leq -C(\log \epsilon_1 + \log(|x-y|)) \leq \frac{|x-y|}{\epsilon_1}. \end{aligned}$$

Similarly one can show that $|H_{\epsilon_2,\epsilon_2}(y,y) - H_{\epsilon_1,\epsilon_2}(x,y)| \leq \frac{|x-y|}{\epsilon_1}$.

Case 3 Let $\overline{D(y,\,\epsilon_2)} \subseteq D(x,\,\epsilon_1)$.

$$\begin{aligned}|H_{\epsilon_1,\epsilon_1}(x,x) - H_{\epsilon_1,\epsilon_2}(x,y)| &\leq |G(\epsilon_1)(1 - I_0(|x-y|))| \\ &\quad + C\frac{I_2(|x-y|)}{I_1^2(\epsilon_1) - I_0(\epsilon_1)I_2(\epsilon_1)} \\ &\leq -C\log\epsilon_1\,|x-y|^2 + \frac{|x-y|^2}{\epsilon_1^2} \leq C\frac{|x-y|}{\epsilon_1}.\end{aligned}$$

Combining these three cases we obtain that

$$\mathsf{Var}_{\mathcal{W}}\left(\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right) - \mathcal{I}\left(h_{\mu^y_{\epsilon_2}}\right)\right) \leq C\frac{|x-y| + |\epsilon_1 - \epsilon_2|}{\epsilon_1 \wedge \epsilon_2}. \tag{2.18}$$

Since $\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right) - \mathcal{I}\left(h_{\mu^y_{\epsilon_2}}\right)$ is Gaussian,

$$\mathsf{E}_{\mathcal{W}}\left[\left|\mathcal{I}\left(h_{\mu^x_{\epsilon_1}}\right) - \mathcal{I}\left(h_{\mu^y_{\epsilon_2}}\right)\right|^{\alpha}\right] \leq C\left(\frac{|x-y| + |\epsilon_1 - \epsilon_2|}{\epsilon_1 \wedge \epsilon_2}\right)^{\alpha/2}.$$

We can find α and β large enough such that $\left|\frac{\beta}{\alpha} - \frac{1}{2}\right| < \delta$, and consequently by [46, Lemma C.1] there exists a modification $\widetilde{X}(x,\epsilon) = \mathcal{I}\left(h_{\mu^x_\epsilon}\right)$ a.s. on $L^2(\mathcal{W})$ satisfying (2.17). $\square$

In this section for the proof of the upper bound we work with this modification which we also denote by $\mathcal{I}\left(h_{\mu^x_t}\right)$. Recall that $B(x,t) = \mathcal{I}\left(h_{\mu^x_{G^{-1}(t)}}\right)$.

Proof of the upper bound. Let $\varepsilon > 0$ and $\gamma \in (0,1/2)$, $\zeta \in (0,1)$ and denote $\tilde{\gamma} := (1+\varepsilon)\gamma$. Also let $K := \varepsilon^{-1}$, $r_n := n^{-K}$.

Define the set

$$U_R := \left\{x \in D(0,R) : \limsup_{n\to+\infty} \frac{\mathcal{I}\left(h_{\mu^x_{r_n}}\right)}{\sqrt{2\pi^2}G(r_n)} \geq \sqrt{2a}\right\}.$$

We first show that

$$T_{\geq}(a,R) \subset U_R. \tag{2.19}$$

For $x \in T_{\geq}(a,R)$ and for $t \in (G(r_n),\,G(r_{n+1}))$ we write

$$B(x,G(r_n)) = B(x,G(r_n)) - B(x,t) + B(x,t).$$

By Proposition 2.12 we have

$$
\begin{aligned}
|B(x,t) - B(x, G(r_n))| &\leq M\left(\log\left(\frac{1}{G^{-1}(t)}\right)\right)^{\zeta} \frac{\left(G^{-1}(t) - r_n\right)^{\gamma}}{G^{-1}(t)^{\tilde{\gamma}}} \\
&\leq M(\log(n+1))^{\zeta} \frac{(r_{n+1} - r_n)^{\gamma}}{r_{n+1}^{\tilde{\gamma}}} = O\left((\log n)^{\zeta}\right). \qquad (2.20)
\end{aligned}
$$

Hence using the fact that $G(r_n) \sim C \log n$ for $n \to +\infty$ and $\zeta < 1$ we have

$$
\left|\frac{B(x, G(r_n)) - B(x,t)}{\sqrt{2\pi^2} G(r_n)}\right| = O\left(\frac{(\log n)^{\zeta}}{G(r_n)}\right) = o\left(1\right).
$$

Now (2.19) follows as we have

$$
\limsup_{n \to +\infty} \frac{B(x, G(r_n))}{\sqrt{2\pi^2} G(r_n)} \geq \limsup_{t \to +\infty} \frac{B(x,t)}{\sqrt{2\pi^2} t} \geq \sqrt{2a}.
$$

The next step is to determine a cover for the set U_R. In view of that, let $(x_{nj})_{j=1}^{\bar{k}_n}$ be a maximal collection of points in $D(0,R)$ such that $\inf_{l \neq j} \left|x_{nj} - x_{nl}\right| \geq r_n^{1+\varepsilon}$. Denote

$$
\mathcal{A}_n := \left\{ j : \frac{\left|B(x_{nj}, G(r_n))\right|}{\sqrt{2\pi^2} G(r_n)} \geq \sqrt{2a} - \delta(n) \right\}
$$

with $\delta(n) = C(\log n)^{\zeta - 1}$ (the constant C will be tuned later according to (2.21)). For any $x \in D(0,R)$, there exists $j \in \left\{1, \ldots, \bar{k}_n\right\}$ such that $x \in D\left(x_{nj}, r_n^{1+\varepsilon}\right)$. By (2.20) we have,

$$
\begin{aligned}
&\frac{\left|B(x_{nj}, G(r_n)) - B(x, G(r_n))\right|}{\sqrt{2\pi^2} G(r_n)} \leq C(\log n)^{\zeta} \frac{\left|x - x_{nj}\right|^{\gamma}}{G(r_n)^{\tilde{\gamma}+1}} \\
&= \delta(n) \frac{\log n}{G(r_n)} \leq C\delta(n) \qquad (2.21)
\end{aligned}
$$

which implies, renaming possibly $\delta(n)$,

$$
\frac{B(x_{nj}, G(r_n))}{G(r_n)} \geq \sqrt{2a} - \delta(n).
$$

Hence we have $j \in \mathcal{A}_n$. Therefore for all $N \geq 1$, $\bigcup_{n \geq N} \bigcup_{j \in \mathcal{A}_N} D\left(x_{nj}, r_n^{1+\varepsilon}\right)$ covers U_R with sets having maximal diameter $2r_n^{1+\varepsilon}$. Next we claim that

$$
\mathsf{E}_{\mathcal{W}}\left[|\mathcal{A}_n|\right] \leq C(\log n) r_n^{a - 4(1+\varepsilon) + \mathrm{o}(1)}. \qquad (2.22)
$$

Assume (2.22) for the moment. If we choose $\alpha := 4 - a + \varepsilon\frac{4+a}{1+\varepsilon}$ we have

$$\begin{aligned}
&\mathsf{E}_{\mathcal{W}}\left[\sum_{n\ge N}\sum_{j\in\mathcal{A}_n}\operatorname{diam}(D(x_{nj}, r_n^{1+\varepsilon}))^{\alpha}\right] \le \sum_{n\ge N}(\log n) r_n^{(1+\varepsilon)\alpha+a-4(1+\varepsilon)+o(1)} \\
&\le \sum_{n\ge N}(\log n) r_n^{4\varepsilon+o(1)} = C\sum_{n\ge N}(\log n) n^{-4+o(1)} < +\infty.
\end{aligned}$$

Therefore $\sum_{n\ge N}\sum_{j\in\mathcal{A}_n}\operatorname{diam}(D(x_{nj}, r_n^{1+\varepsilon}))^{\alpha} < +\infty$ a. s. and this implies that $\dim_{\mathcal{H}}(T_{\ge}(a,r)) \le 4 - a$ a.s. by letting $\varepsilon \downarrow 0$. This completes the proof of the upper bound provided we show (2.22). We first estimate $\mathcal{W}\,(j \in \mathcal{A}_n)$ as follows:

$$\begin{aligned}
&\mathcal{W}\,(j\in\mathcal{A}_n) = \mathcal{W}\left(\frac{|B(x_{nj}, G(r_n))|}{\sqrt{G(r_n)}} \ge (\sqrt{2a} - \delta(n))\sqrt{2\pi^2}\sqrt{G(r_n)}\right) \\
&\le C(a + o\,(1))G(r_n)\exp\left(-a\,(1-o\,(1))^2\,2\pi^2 G(r_n)\right) \le C(\log n) r_n^{a+o(1)},
\end{aligned}$$

since $G(r_n) \sim -\frac{\log r_n}{2\pi^2}$ as $n \to +\infty$. Furthermore

$$\mathsf{E}_{\mathcal{W}}\left[|\mathcal{A}_n|\right] \le C(\log n)\bar{k}_n r_n^{(a+o(1))} \le (\log n) r_n^{a+o(1)-4(1+\varepsilon)}.$$

This proves (2.22) and hence the upper bound.

Now we show that for every $R > 1$, $T_{\ge}(a,R)$ is empty for $a > 4$ using the above estimates. Note that

$$\sum_{n\ge 1}\mathcal{W}\,(|\mathcal{A}_n| > 1) \le \sum_{n\ge 1}\mathsf{E}_{\mathcal{W}}\left[|\mathcal{A}_n|\right] \le \sum_{n\ge 1} r_n^{a-4(1+\varepsilon)} = \sum_{n\ge 1} r_n^4 < +\infty$$

and hence by the Borel-Cantelli lemma we can conclude that, if ε becomes arbitrarily small, $|\mathcal{A}_n| = 0$ eventually and so $T_{\ge}(a,R)$ is empty for $a > 4$. □

2.3.7. Lower bound of Theorem 2.8

To derive the lower bound we use the energy method. For detailed use of this method see Section 4.3 of [64]. The α-th energy of a measure μ is given by

$$I_\alpha(\mu) = \iint \frac{\mathrm{d}\mu(x)\mathrm{d}\mu(y)}{|x-y|^{\alpha}}.$$

Given a set A, if we can find a measure ρ such that $I_\alpha(\rho) < \infty$ then $\dim_H(A) > \alpha$. For this, partition the hypercube $J := [0,1]^4$ into s_n^{-4} smaller hypercubes of radius

$s_n = \frac{1}{n!}$. Let x_{ni} be the centers of these hypercubes and C_n be the set of these centers. Define $t_m := -\frac{\log s_m}{2\pi^2}$ for all $m \leq n$ and let $A_m(x)$, $B_m(x)$ be the events

$$A_m(x) := \left\{ \sup_{t_m < t \leq t_{m+1}} \left| B(x,t) - B(x,t_m) - \sqrt{4a\pi^2}(t - t_m) \right| \leq \sqrt{t_{m+1} - t_m} \right\},$$

$$B_m(x) := \left\{ \sup_{t \geq t_m} |B(x,t) - B(x,t_m)| - t \leq 1 - t_m \right\}$$

We say that x is an *n-perfect a-thick point* if $E^n(x) := \bigcap_{m \leq n} A_m(x) \cap B_{n+1}(x)$ occurs. Note that $B_{n+1}(x)$ is independent of the other events. We introduce a random variable Y_{ni} for $i = 1, \cdots, |C_n|$ such that

$$Y_{ni} = \begin{cases} 1 & \text{if } x_{ni} \text{ is an } n\text{-perfect } a\text{-thick point,} \\ 0 & \text{otherwise.} \end{cases}$$

Fix $t_m < t \leq t_{m+1}$ and on the event $E^n(x)$ we have, as $m \to \infty$,

$$\left| B(x,t) - B(x,t_1) - \sqrt{4a\pi^2}(t - t_1) \right| = o\left(m \log m\right) = o\left(t\right). \tag{2.23}$$

Define now the set of perfect a-thick points as

$$P(a) := \bigcap_{k \geq 1} \overline{\bigcup_{n \geq k} \bigcup_{z \in C_n(a)} S(z, s_n)},$$

where $C_n(a)$ is the set of centers of which x_{ni} is a n-perfect thick point and $S(z,r)$ is a hypercube of radius r centered around z. Let

$$T(a,J) := \left\{ x \in J : \lim_{t \to \infty} \frac{\mathcal{I}\left(h_{\mu^x_{G^{-1}(t)}}\right)}{\sqrt{2\pi^2 t}} = a \right\} \subset T(a).$$

Lemma 2.13.

$$P(a) \subseteq T(a,J). \tag{2.24}$$

Proof. If $z \in P(a)$ there exists a sequence $(z_{n_k})_{k \in \mathbb{N}}$ of points s. t. $z_{n_k} \in C_n(a)$ for all k and $|z - z_{n_k}| \leq s_n$. For m s. t. $t_m < t \leq t_{m+1}$

$$\left| B(z_{n_k}, t) - B(z_{n_k}, t_1) - \sqrt{4a\pi^2}(t - t_1) \right| = o\left(t\right)$$

follows as in (2.23). Since the Brownian motion is a.s. continuous taking the limit for $k \to +\infty$

$$\left| B(z,t) - B(z,t_1) - \sqrt{4a\pi^2}(t - t_1) \right| = o\left(t\right)$$

and dividing by $\sqrt{2\pi^2}t$

$$\left| \frac{\mathcal{I}\left(h_{\mu^z_{G^{-1}(t)}}\right)}{\sqrt{2\pi^2}t} - \sqrt{2a} \right| = o\left(1\right)$$

which is an equivalent formulation of the set of thick points. □

Next we make preparations to define a measure μ supported on $P(a)$ with positive probability. For this purpose define a sequence of measures μ_n on J supported on n-perfect thick points.

$$\mu_n(\cdot) = \sum_{i=1}^{|C_n|} \frac{1}{\mathcal{W}\left(E^n(x_{ni})\right)} \mathbb{1}_{\{Y_{ni}=1\}} \lambda\left(\cdot \cap S(x_{ni}, s_n)\right), \tag{2.25}$$

where $\lambda(\cdot)$ is the Lebesgue measure.

In the following lemma we list down some important properties of this measure.

Lemma 2.14. *Let $\mu_n(\cdot)$ be as above. Then the following hold:*

(a) $\mathsf{E}_{\mathcal{W}}\left[\mu_n(J)\right] = 1$;

(b) $\sup_n \mathsf{E}_{\mathcal{W}}\left[\mu_n(J)^2\right] < \infty$;

(c) $\sup_n \mathsf{E}_{\mathcal{W}}\left[I_\alpha(\mu_n)\right] < \infty$;

(d) there exist $a, b \in (0, \infty)$ such that for all n we have

$$\mathcal{W}\left(b \le \mu_n(J) < b^{-1}, I_\alpha(\mu_n) < a\right) > 0.$$

The proof of Lemma 2.14 requires a correlation inequality and a lower bound depends on the following lemma. Its proof is similar to the proof of Lemma 3.3 of [46] and hence we briefly sketch it.

Lemma 2.15. *Let $A_m(x), B_m(x)$ be as above with $s_m = \frac{1}{m!}$. Let*

$$E^n(x) = \bigcap_{m \le n} A_m(x) \cap B_{n+1}(x).$$

Then for every $y \in S(x, s_l) \setminus S(x, s_{l+1})$, $l > 2$, we have

$$\mathcal{W}\left(E^n(x) \cap E^n(y)\right) \le \mathscr{C}_l \mathcal{W}\left(E^n(x)\right) \mathcal{W}\left(E^n(y)\right), \tag{2.26}$$

where $\mathscr{C}_l$ is defined by

$$\mathscr{C}_l := C \prod_{j \le l+1} j^a \exp\left(-\frac{a}{2}\sqrt{\log j}\right).$$

Proof of Lemma 2.15. Fix $l > 2$ and $y \in S(x, s_l) \setminus S(x, s_{l+1})$. First note that the collections $\{A_i(x) : 1 \le i \le l+1\}$ and $\{A_i(x) : l+2 \le i \le n\}$ are independent as they depend on disjoint annuli. Similarly, as $S(x, s_{l+2}) \cap S(x, s_j) \setminus S(x, s_{j+1}) = \emptyset$ the collection $\{A_j(y) : j \ne l-1, l, l+1\}$ and $\{A_i(x) : l+2 \le i \le n\}$ are independent. Now note that by the assumption,

$$\begin{aligned}
&\mathcal{W}\left(\bigcap_{1\le i\le l+1} A_i(x)\right)\mathcal{W}\left(\bigcap_{l-1\le j\le l+1} A_j(y)\right) \\
&= \prod_{1\le i\le l+1} \mathcal{W}\left(A_i(x)\right) \prod_{l-1\le j\le l+1} \mathcal{W}\left(A_j(y)\right) \ge \prod_{i=1}^{l+1} \mathscr{C}_l. \qquad (2.27)
\end{aligned}$$

So we have

$$\begin{aligned}
\mathcal{W}\left(E^n(x) \cap E^n(y)\right) &= \mathcal{W}\left(\bigcap_{i\le n} A_i(x) \cap B_{n+1}(x) \cap \bigcap_{j\le n} A_j(y) \cap B_{n+1}(y)\right) \\
&\le \mathcal{W}\left(\bigcap_{i\le n} A_i(x) \cap \bigcap_{j\le n} A_j(y)\right) \\
&\le \mathcal{W}\left(\bigcap_{l+2\le i\le n} A_i(x) \cap \bigcap_{j\le n, j\ne l-1,l,l+1} A_j(y)\right) \\
&\le \mathcal{W}\left(\bigcap_{l+2\le i\le n} A_i(x)\right)\mathcal{W}\left(\bigcap_{j\le n, j\ne l-1,l,l+1} A_j(y)\right)
\end{aligned}$$

If we now multiply and divide the last probability by

$$\mathcal{W}\left(\bigcap_{1\le i\le l+1} A_i(x)\right)\mathcal{W}\left(\bigcap_{l-1\le j\le l+1} A_j(y)\right)$$

and use independence we get

$$\mathcal{W}\left(E^n(x) \cap E^n(y)\right) \le \frac{\mathcal{W}\left(\bigcap_{i\le n} A_i(x)\right)\mathcal{W}\left(\bigcap_{i\le n} A_i(y)\right)}{\mathcal{W}\left(\bigcap_{1\le i\le l+1} A_i(x)\right)\mathcal{W}\left(\bigcap_{l-1\le j\le l+1} A_j(y)\right)}.$$

Now using the bound in (2.27) and the fact that $B_{n+1}(x)$ is independent from $\{A_i(x) : i \le n\}$ we get

$$\mathcal{W}\left(E^n(x) \cap E^n(y)\right) \le \mathscr{C}_l \mathcal{W}\left(E^n(x)\right)\mathcal{W}\left(E^n(y)\right)$$

We can adjust appropriately the constant $\mathscr{C}_l$ when $l \le 2$ to complete the proof. $\square$

Using the above Lemma the proof of Lemma 2.14 is almost immediate.

Proof of Lemma 2.14. Note the series $\sum_{l=1}^{\infty} s_l^4 \mathscr{C}_l$ converges (absolutely) by the ratio test. By means of the same criterion one shows also that $\sum_{l=1}^{\infty} s_l^4 \mathscr{C}_l s_{l+1}^{-\alpha} < +\infty$ under the assumption $\alpha \leq 4$. Keeping these facts in mind we proceed to the proof.

(a) As $S(x_{ni}, s_n)$ forms a cover of J it is easy to show that $\mathsf{E}_{\mathcal{W}}[\mu_n(J)] = 1$. In particular,

$$\begin{aligned}\mathsf{E}_{\mathcal{W}}[\mu_n(J)] &= \sum_{i=1}^{|C_n|} \frac{1}{\mathcal{W}(E^n(x_{ni}))} \mathcal{W}(Y_{ni} = 1)\, \lambda(J \cap S(x_{ni}, s_n)) \\ &= \sum_{i=1}^{|C_n|} \lambda(J \cap S(x_{ni}, s_n)) = 1.\end{aligned}$$

(b) Using Lemma 2.15 we have,

$$\begin{aligned}\mathsf{E}_{\mathcal{W}}\left[\mu_n(J)^2\right] &= \sum_{i,j=1}^{|C_n|} \frac{\mathcal{W}\left(Y_{ni} = 1, Y_{nj} = 1\right)}{\mathcal{W}(E^n(x_{ni}))\,\mathcal{W}\left(E^n(x_{nj})\right)} \lambda(S(x_{ni}, s_n))\lambda(S(x_{nj}, s_n)) \\ &\leq s_n^8 \sum_{i=1}^{|C_n|} \sum_{l=1}^{n} \sum_{j=1, s_l \geq |x_{nj} - x_{ni}| > s_{l+1}}^{n} \frac{\mathcal{W}\left(E^n(x_{ni}) \cap E^n(x_{nj})\right)}{\mathcal{W}(E^n(x_{ni}))\,\mathcal{W}\left(E^n(x_{nj})\right)} \\ &\leq s_n^8 \sum_{i=1}^{|C_n|} \sum_{l=1}^{n} \frac{s_l^4}{s_n^4} \mathscr{C}_l \leq \sum_{l=1}^{\infty} s_l^4 \mathscr{C}_l < \infty.\end{aligned}$$

Above we have used the fact that the number of hypercubes with center at x_{ni} and radius s_l is proportional to s_l^4 / s_n^4.

(c) For the expected energy we follow the same procedure as above. Note that $|x_{ni} - x_{nj}| > s_{l+1}$ then if we take $x \in S(x_{ni}, s_n)$ and $y \in S(x_{nj}, s_n)$ then $|x - y| > s_{l+1}$.

$$\begin{aligned}\mathsf{E}_{\mathcal{W}}[I_\alpha(\mu_n)] &= \sum_{i,j=1}^{|C_n|} \frac{\mathcal{W}\left(E^n(x_{ni}) \cap E^n(x_{nj})\right)}{\mathcal{W}(E^n(x_{ni}))\,\mathcal{W}\left(E^n(x_{nj})\right)} \int_{S(x_{ni}, s_n)} \int_{S(x_{nj}, s_n)} \frac{\mathrm{d}x\mathrm{d}y}{|x - y|^\alpha} \\ &\leq s_n^8 \sum_{i=1}^{|C_n|} \sum_{l=1}^{n} \frac{s_l^4}{s_n^4} \mathscr{C}_l s_{l+1}^{-\alpha} \leq \sum_{l \geq 1} \mathscr{C}_l s_l^4 s_{l+1}^{-\alpha} < +\infty.\end{aligned}$$

(d) By Paley-Zygmund inequality we have

$$\mathcal{W}\left(b \leq \mu_n(J) \leq b^{-1}\right) \geq (1-b)^2 \frac{1}{\mathsf{E}_{\mathcal{W}}[\mu_n(J)^2]} - (1-b)^2 \frac{1}{\mathsf{E}_{\mathcal{W}}[\mu_n(J)^2]}.$$

Now using part (b) we have that for all n, $\mathcal{W}\left(b \le \mu_n(J) \le b^{-1}\right) > 0$. Also note that since $\mathsf{E}_{\mathcal{W}}\left[I_\alpha(\mu_n)\right]$ is uniformly bounded,

$$\mathcal{W}\left(I_\alpha(\mu_n) \le a\right) > 0.$$

Hence (d) follows from the above observations.

$\square$

Proof of the lower bound. Now using Lemma 2.14 we continue with the proof of lower bound. If we define

$$G := \limsup_{n\to+\infty} \left\{b \le \mu_n(J) < b^{-1}, I_\alpha(\mu_n) < a\right\},$$

then by Lemma 2.14 (d), $\mathcal{W}(G)$ is bounded away from zero. I_α being a lower semi-continuous function, the set of measures μ for which $b \le \mu(J) < b^{-1}$ and $I_\alpha(\mu) < a$ is compact in the topology of weak convergence. Therefore the sequence $(\mu_n)_{n\in\mathbb{N}}$ admits surely along a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ a weak limit μ, which is a finite measure supported on $P(a)$ and whose α-energy is finite. Hence, we have

$$\mathcal{W}\left(C_{\mathcal{H}}^{4-a}(P(a)) > 0\right) > 0. \tag{2.28}$$

Now by the monotonicity of the Hausdorff-α-measure, if we can show that

$$\mathcal{W}\left(C_{\mathcal{H}}^{4-a}(T(a,J)) > 0\right) \in \{0,1\}$$

then by (2.28), the set $\left\{C_{\mathcal{H}}^{4-a}(T(a,J)) > 0\right\}$ will have probability one and hence the proof will be complete.

Now from the construction of μ_ϵ^x, it holds from Equation 7.9 of [20] that $\mathcal{I}\left(h_{\mu_\epsilon^x}\right) = f_1(\epsilon)\mathcal{I}(h_{\sigma_\epsilon^x}) + f_2(\epsilon)\mathcal{I}(h_{\mathrm{d}\sigma_\epsilon^x})$, where

$$f_1(\epsilon) = \frac{\epsilon I_1(\epsilon) - 2I_2(\epsilon)}{I_1^2(\epsilon) - I_0(\epsilon)I_2(\epsilon)}, \quad f_2(\epsilon) = \frac{-\epsilon I_2(\epsilon)}{I_1^2(\epsilon) - I_0(\epsilon)I_2(\epsilon)}.$$

Since $\lim_{\epsilon\to 0} f_1(\epsilon) = 2$ and $\lim_{\epsilon\to 0} f_2(\epsilon) = 0$, $\mu_\epsilon^x \to 2\delta_x$ as $\epsilon \to 0$ in the sense of distributions. In fact, since $\widehat{\mathrm{d}\sigma_\epsilon^x}(\xi) = -\frac{2}{\epsilon}J_2(\epsilon|\xi|)\exp\left(i(\xi,x)_{\mathbb{R}^4}\right) \to 0$ for all ξ, $\mathrm{d}\sigma_\epsilon^x \to 0$ in the sense of distributions. Thus

$$\limsup_{\epsilon\to 0} \frac{\mathcal{I}\left(h_{\mu_\epsilon^x}\right)}{\sqrt{2\pi}G(\epsilon)} = \limsup_{\epsilon\to 0} \frac{f_1(\epsilon)\mathcal{I}(h_{\sigma_\epsilon^x})}{\sqrt{2\pi}G(\epsilon)}.$$

By [74] (Section 2), if $\{h_m\}$ is an orthonormal basis of H,

$$[\mathcal{I}(h_{\sigma_\epsilon^x})](\theta) = \langle \theta, \sigma_\epsilon^x \rangle \overset{\mathcal{W}-\text{a.s.}}{=} \left\langle \sum_{m \geq 1} [\mathcal{I}(h_m)(\theta)] h_m, \sigma_\epsilon^x \right\rangle .$$

The series will depend then only on its tail, as $\langle h_m, \sigma_\epsilon^x \rangle \to h_m(x)$ and $G(\epsilon) \to +\infty$ as $\epsilon \to 0$. Using the fact that $(\mathcal{I}(h_m))_{m \geq 1}$ are i.i.d. we can apply Kolmogorov's 0-1 law to conclude. □

Part II.

Random permutations

Chapter 3.

Random permutations

A *random permutation* on n objects is a random shuffling of those objects, viz. a permutation-valued random variable. Their applications are extremely wide, ranging from cryptography and computer science to number theory and genetics. For example, the statistics of random permutations, such as the cycle structure of a random permutation, are of fundamental importance in the analysis of algorithms, especially of sorting ones which act on "disordered" sets. This is were probabilty kicks in, as it allows us to calculate the expected behavior of random orderings in order to control them.

3.0.8. The Young diagram

We denote by $\mathfrak{S}_n$ the set of permutations on n elements.

Definition 3.1. *A* cycle $\sigma := (s_0\, s_1 \ldots s_{k-1})$ *is a permutation belonging to* $\mathfrak{S}_n$ *which maps*

$$s_0 \mapsto s_1 \mapsto s_2 \mapsto \cdots \mapsto s_{k-1} \mapsto s_k = s_0.$$

and agrees with the identity on the remaining points. Two cycles $(s_0 \ldots s_{k-1})$ *and* $(t_0 \ldots t_{m-1})$ *are called disjoint if the sets* $\{s_0, \ldots, s_{k-1}\}$ *and* $\{t_0, \ldots, t_{m-1}\}$ *are disjoint.*

If $\sigma \in \mathfrak{S}_n$ is given, then ([5]) it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles. We write λ_j for the length of the cycle σ_j. W. l. o. g. we can assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$. The partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is called the *cycle type* of σ and we will indicate it by $\lambda \vdash n$. λ features a nice geometric visualisation by its Young diagram Y_λ. This is a left- and bottom-justified diagram of ℓ rows with the $j-$th row consisting of λ_j squares. It is clear that the area of Y_λ is n if $\lambda \vdash n$. The upper boundary of a Young diagram Y_λ is a piecewise constant and right continuous function $w_n : \mathbb{R}^+ \to \mathbb{N}^+$ with

$$w_n(x) := \sum_{j=1}^{n} \mathbb{1}_{\{\lambda_j \geq x\}}\,. \tag{3.0.1}$$

The notation $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is useful for the illustration via Young diagrams. For computations, we prefer the notation involving cycle counts.

Definition 3.2. *Given a permutation $\sigma \in \mathfrak{S}_n$ we write*

$$C_k^{(n)} = C_k^{(n)}(\sigma) := |\{j : \lambda_j = k\}|$$

for all $1 \leq k \leq n$.

We can write formally

$$\lambda = \left(1^{C_1^{(n)}}, 2^{C_2^{(n)}}, \ldots\right).$$

meaning that the partition λ consists in $C_i^{(n)}$ block of length i for any $i \leq n$. This then gives

$$w_n(x) := \sum_{k \geq x}^{n} C_k^{(n)}.$$

An example of Young diagram can be seen in Figure 3.0.8. In this paper, we work with the following measure on $\mathfrak{S}_n$:

$$\mathbb{P}_n[\sigma] = \frac{1}{h_n n!} \prod_{j=1}^{\ell} \vartheta_{\lambda_j}. \tag{3.0.2}$$

where $(\lambda_1, \ldots, \lambda_\ell)$ is the cycle type of σ, $(\vartheta_m)_{m \geq 1}$ is a sequence of non-negative weights and h_n is a normalization constant (h_0 is defined to be 1). From time to time we will also use $\vartheta_0 := 0$ introduced as convention. $\mathbb{P}_n[\,\cdot\,]$ in (3.0.2) can now be written as

$$\mathbb{P}_n[\sigma] = \frac{1}{h_n n!} \prod_{k=1}^{n} \vartheta_k^{C_k} \tag{3.0.3}$$

with the convention that $C_0^{(n)} := 0$. One has

Lemma 3.3. *Let $c_1, c_2, \ldots, c_n \in \mathbb{N}$ be given. Then*

$$\mathbb{P}_n\left(C_1^{(n)} = c_1, \ldots, C_n^{(n)} = c_n\right) = \prod_{k_1}^{n} \left(\frac{1}{k}\right)^k \frac{1}{c_k} \mathbb{1}_{\left\{\sum_{k \leq n} k c_k = n\right\}}.$$

Lemma 3.4 ([32], Corollary 2.3)**.** *Under the condition $h_{n-1}/h_n \to 1$ the random variables C_k converge for each $k \in \mathbb{N}$ in distribution to a Poisson distributed random variable Y_k with $\mathbb{E}[Y_k] = \frac{\vartheta_k}{k}$. More generally for all $b \in \mathbb{N}$ the following limit in distribution holds:*

$$\lim_{n \to +\infty} (C_1, C_2 \ldots, C_b) = (Y_1, Y_2 \ldots, Y_b)$$

with Y_k independent Poisson random variables with mean $\mathbb{E}[Y_k] = \frac{\vartheta_k}{k}$.

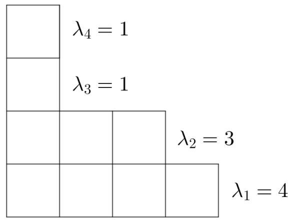

Figure 3.1.: Young diagram for $\sigma = (128)(3)(4579)(6) \in \mathfrak{S}_9$ with $\lambda = (4,3,1,1)$.

We will prove the corresponding version of the above lemmas in our case in Lemma 3.8. What we are then interested in studying is the now random shape $w_n(\cdot)$ as $n \to +\infty$, and more specifically to determine its limit shape. The limit shape with respect to a sequence of probability measures $\mathbb{P}_n$ on $\mathfrak{S}_n$ (and sequences of positive real numbers A_n and B_n with $A_n \cdot B_n = n$) is understood as a function $w_\infty : \mathbb{R}^+ \to \mathbb{R}^+$ such that for each $\epsilon, \delta > 0$

$$\lim_{n\to+\infty} \mathbb{P}_n \left[\left\{ \sigma \in \mathfrak{S}_n : \sup_{x\geq\delta} |A_n^{-1} w_n(B_n x) - w_\infty(x)| \leq \epsilon \right\} \right] = 1 \tag{3.0.4}$$

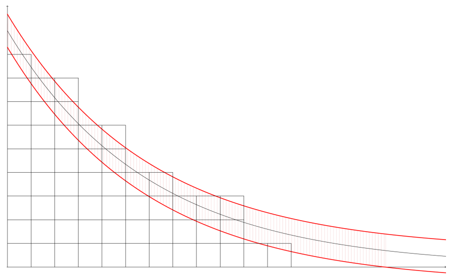

Figure 3.2.: Visualization of the limit shape.

The assumption $A_n \cdot B_n = n$ ensures that the area under the rescaled Young diagram is 1. One of the most frequent choices is $A_n = B_n = n^{1/2}$, but we will see that it's useful to adjust the choice of A_n and B_n to the measures $\mathbb{P}_n$. Equation 3.0.4 can be viewed as a law of large numbers for the process $w_n(\cdot)$. This measure has recently appeared in mathematical physics for a model of the quantum gas in statistical mechanics and has a possible connection with the Bose-Einstein condensation (see e.g. [10] and [32]). Classical cases of this measure are the uniform measure ($\vartheta_m \equiv 1$) and the Ewens measure ($\vartheta_m \equiv \vartheta$). The uniform measure is well studied and has a long history (see

e.g. the first chapter of [4] for a detailed account with references). The Ewens measure originally appeared in population genetics, see [35], but has also various practical applications through its connection with Kingman's coalescent process, see [45]. The measure in (3.0.2) also has some similarities to multiplicative measure for partitions, see for instance [13].
It is clear that we have to make some assumptions on the sequence $(\vartheta_m)_{m\geq 1}$ to be able study the behaviour as $n \to +\infty$. We use in this paper cycle weights ϑ_m considered in the recent work Ercolani and Ueltschi [32] and of Maples, Nikeghbali and Zeindler [63]. In view of the argumentation, it is the most natural to specify this assumption using the function

$$g_\Theta(t) = \sum_{m\geq 1} \frac{\vartheta_m}{m} t^m \tag{3.0.5}$$

We may mention two popular methods in the literature to study the asymptotic behaviour of the function $w_n(x)$ under such assumptions. The first one is is complex analytic and uses the saddle-point method as described in Section 3.3, whilst the second one is stochastic and based on randomisation, and we will present it in Section 3.2. It is typically expected that the 'randomised' $w_n(x)$ has the same asymptotic behaviour as the 'unrandomized' $w_n(x)$. We will see here that this is not the case. More precisely, we show that the 'randomized' and the 'unrandomized' $w_n(x)$ have the same limit shape but different behaviors of the fluctuations around the limit shape. Finally for the 'unrandomized' case we will prove a functional CLT for the limit shape.

$$C_k \geq 0 \ \text{ and } \ \sum_{k=1}^{n} kC_k = n. \tag{3.0.6}$$

It is also clear that the cycle type of permutation (or a partition) is uniquely determined by the vector $(C_1, C_2, \dots)$.

Our aim is to study the behaviour of $w_n(x)$ as $n \to \infty$. It is thus natural to consider the asymptotic behaviour of C_k with respect to the measure $\mathbb{P}_n[\,\cdot\,]$.

One might expect at this point that $w_n(x)$ is close to $\sum_{k=x}^{n} Y_k$, we will see though in Section 3.3 that the asymptotic behaviour of $w_n(x)$ is more complicate.

Generating functions

The (ordinary) generating function of a sequence $(a_k)_{k\geq 0}$ of complex numbers is defined as the formal power series

$$g(z) := \sum_{j=0}^{\infty} a_k z^k. \tag{3.0.7}$$

As usual, we define the *extraction symbol* $[z^k]\, g(z) := a_k$, that is, as the coefficient of z^k in the power series expansion (3.0.7) of $g(z)$.

A generating function that plays an important role in this paper is $g_\Theta(t)$ as in (3.0.5) where

$$\vartheta_m = \frac{m^\alpha}{\Gamma(\alpha+1)} + O\left(m^\beta\right) \tag{3.0.8}$$

and $0 \le \beta < \alpha$. We stress that generating functions of the type $(1-t)^{-\alpha}$ fall also in this category, and for them we will recover the limiting shape as previously done in [34]. We will see in particular this case in Section 3.3.
The reason why generating functions are useful is that it is often possible to write down a generating function without knowing a_n explicitly. In this case one can try to use tools from analysis to extract information about a_n, for large n, from the generating function. It should be noted that there are several variants in the definition of generating functions. However, we will use only the ordinary generating function and thus call it 'just' generating function without risk of confusion.

The following well-known identity is a special case of the general *Pólya's Enumeration Theorem* [68] and is the main tool in this paper to obtain generating functions.

Lemma 3.5. *Let $(a_m)_{m\in\mathbb{N}}$ be a sequence of complex numbers. We then have as formal power series in t*

$$\sum_{n\in\mathbb{N}} \frac{t^n}{n!} \sum_{\sigma\in\mathfrak{S}_n} \prod_{j=1}^{n} a_j^{C_j} = \sum_{n\in\mathbb{N}} t^n \sum_{\lambda\vdash n} \frac{1}{z_\lambda} \prod_{m\ge 1} \prod_{k=1}^{\infty} a_k^{C_k} = \exp\left(\sum_{m\ge 1} \frac{a_m}{m} t^m\right)$$

where $z_\lambda := \prod_{k=1}^{n} k^{C_k} C_k!$. If one series converges absolutely, so do the others.

We omit the proof of this lemma, but details can be found for instance in [61, p. 5].

Approximation of sums

We require for our argumentation the asymptotic behaviour of the generating function $g_\Theta(t)$ as t tends to the radius of convergence, which is 1 in our case.

Lemma 3.6. *Let $(v_n)_{n\in\mathbb{N}}$ a sequence of postive numbers with $v_n \downarrow 0$ as $n \to +\infty$. We have for all $\delta \in \mathbb{R} \setminus \{-1, -2, -3, \dots\}$*

$$\sum_{k=1}^{\infty} k^\delta e^{-kv_n} = \Gamma(\delta+1) v_n^{-\delta-1} + \zeta(-\delta) + O(v_n). \tag{3.0.9}$$

$\zeta(\cdot)$ indicates the Riemann Zeta function.

This lemma can proven with Euler Maclaurin summation formula or with the Mellin transformation. The computations with Euler Maclaurin summation are straightforward and the details of the proof with the Mellin transformation can be found for instance in [36, Chapter VI.8]. We thus omit the proof.

We require also the behavior of the partial sum $\sum_{k=x}^{\infty} \frac{\theta_m}{m} t^m$ as $x \to \infty$ and as $t \to 1$. We have

Lemma 3.7 (Approximation of sums)**.** *Let v_n, z_n be given with $z_n \to +\infty$ and $z_n v_n = a_0 + a_1 n^{-\beta}$ for $\beta > 0, a_0 > 0$ and $a_0, a_1 \in \mathbb{R}$, $a_0 \neq 0$. We then have for all $\delta \in \mathbb{R}$ and all $q \in \mathbb{N}$*

$$\sum_{k=\lfloor z_n \rfloor}^{\infty} k^{\delta} e^{-kv_n} = \left(\frac{z_n}{a_0}\right)^{\delta+1} \left(\sum_{k=0}^{q} \frac{\Gamma(\delta+k+1, a_0)}{k!} \left(-\frac{a_1}{a_0} n^{-\beta}\right)^k + O\left(n^{-(q+1)\beta}\right)\right)$$
$$- B_1(z_n - \lfloor z_n \rfloor) f(z_n) + \int_{z_n}^{+\infty} B_1(y - \lfloor y \rfloor)(\delta - v_n y) y^{\delta-1} e^{-v_n y}\, \mathrm{d}y.$$

with $\Gamma(a,x) := \int_x^{+\infty} s^{a-1} e^{-s} \mathrm{d}s$ the incomplete Gamma function and $B_1(x) := x - \frac{1}{2}$ the first Bernoulli polynomial.

Proof. The proof of this lemma is based on the Euler Maclaurin summation formula, see [3] or [2, Theorem 3.1]. We use the here the following version: let $f : \mathbb{R}^+ \to \mathbb{R}$ have a continuous derivative and suppose that f and f' are integrable. Then

$$\sum_{k \geq \lfloor c \rfloor} f(k) = \int_c^{+\infty} f(x)\, \mathrm{d}x - B_1(c - \lfloor c \rfloor) f(c) + \int_c^{+\infty} B_1(x - \lfloor x \rfloor) f'(x)\, \mathrm{d}x. \tag{3.0.10}$$

A proof of (3.0.10) together with a Euler Maclaurin summation formula with non-integer boundaries can be found in Appendix A.3. We substitute $f(x) := x^{\delta} e^{-xv_n}$, $c := z_n$ and notice that f and all its derivatives tend to zero exponentially fast as $x \to +\infty$. Now by the change of variables $x := \frac{z_n}{a_0} y$

$$\int_{z_n}^{+\infty} e^{-v_n x} x^{\delta} \mathrm{d}x = \left(\frac{z_n}{a_0}\right)^{\delta+1} \int_{a_0}^{+\infty} y^{\delta} e^{-y} e^{-\frac{a_1}{a_0} n^{-\beta} y} \mathrm{d}y =$$
$$= \left(\frac{z_n}{a_0}\right)^{\delta+1} \left(\sum_{k=0}^{q} \frac{\Gamma(\delta+k+1, a_0)}{k!} \left(-\frac{a_1}{a_0} n^{-\beta}\right)^k + O\left(n^{-(q+1)\beta}\right)\right) \tag{3.0.11}$$

where we have swapped integral and series expansion of the exponential by Fubini's theorem. □

3.0.9. A brief historic overview

The evolution of shapes of random ensembles of particles, as the number of particles goes to infinity, was studied for a long time in a variety of applied fields: statistical mechanics (the Wulf construction for the formation of crystals, see [38,39]), stochastic processes on lattices (the Richardson model, see [31]), biology (growth of colonies), etc. A special study was concentrated on limit shapes for random structures on the set of partitions, in view of applications to statistical mechanics, combinatorics, representation theory, and additive number systems. In 1977, two independent teams of researchers, Vershik and Kerov [53] and Shepp and Logan [60], derived the limit shape of a Young diagram with respect to the Plancherel measure. Following this seminal result, Pittel [67] found the limit shape of Young tableaux with respect to a uniform measure. Since the number of Young tableaux corresponding to a given partition (Young diagram) is known to be equal to the degree of the irreducible representation associated with the partition, the above uniform measure, as well as the Plancherel measure, is related to the hook formula, which calculates precisely this quantity. The *Plancherel measure* gives to a partition λ a weight proportional to $\frac{\dim(\lambda)}{n!}$; the "dimension" $\dim(\lambda)$ of a partition is defined as the number of Young tableaux in Young diagrams. It should be mentioned that research concerned with the Plancherel measure revealed also a deep linkage to the random matrix theory, which is now a rapidly growing subject. Parallel to this line of research, Vershik [81] developed a general theory of limit shapes for a class of measures he called multiplicative. These measures encompass a wide scope of models from statistical mechanics and combinatorics, but do not include the measures associated with the hook formula. The results on limit shapes of multiplicative measures obtained by Vershik and Yakubovich [80, 84, 37, 82] during the last decade concern measures induced by Euler type generating functions. Extending these results, Bogachev [13] derived limit shapes for multiplicative measures corresponding to a wide class of generating functions. Note that the limit shape of the uniform measure on the set of partitions (which is a multiplicative measure) was [U+FB01]rstly obtained via a heuristic argument, by Temperley [78]. A comprehensive study of this case was done by Pittel in [66]. Of course this depends heavily on the distribution we set: for example, the most natural choice is the uniform distribution which has a compelling history (see [76], [24], [42],[33]).

3.0.10. Multiplicative measures

Multiplicative measures can be expressed on cycle types in the form

$$\mathbb{P}[\lambda] = \frac{1}{K_n} \prod_{k=1}^{n} b_{C_k}$$

for normalizing constants K_n and weights $b_{C_k} \geq 0$. Their applications are broad and we would like to give here a quick glimpse of four different frameworks in which they have been applied.

Coagulation–fragmentation processes (CFP's) Given an integer N, a CFP is a continuous-time Markov chain on the set Ω_N of all partitions of N. N stands for the total population of indistinguishable particles partitioned into λ_j groups of size j, $j = 1, \ldots, N$. The possible infinitesimal (in time) transitions are coagulation of two groups (=clusters) of size i and j into one group of size $i + j$ and fragmentation of one group of size $i + j$ into two groups of size i and j. If the ratio of these rates is of the form

$$\frac{a_{i+j}}{a_i a_j}$$

$2 \leq i + j \leq N$, $a_i \geq 0$ for all i, then the process is the model formulated in the 1970s by Kelly and Whittle. In particular it can be shown (references can be found in [30, 51]) that the invariant distribution given the parameter a_i is a multiplicative measure.

Decomposable random combinatorial structures A decomposable structure of size N is a union of indecomposable components, so that the counts $n_1, \ldots, n_N$ of components of sizes $1, \ldots, N$ respectively form an integer partition of N. Given a sequence of integers $m = \{m_k\}$, it is assumed that each component of size k belongs to one of the m_k types. Three classes of decomposable structures exist: assemblies, multisets and selections. Supposing that a structure is chosen randomly from the finite set of a certain class of structures with size N, the random partition $\mathbf{K}(N)$ of an integer N is induced:

$$\mathbf{K}(N) = (K_1(N), \ldots, K_N(N)) = \sum_{k=1}^{N} k K_k(N),\ N \geq 1,$$

where the random variable $K_k(N)$ represents the number of components of size k in the random structure. The distributions of $\mathbf{K}(N)$ corresponding to assemblies, multisets and selections are just the multiplicative measures induced by Poisson, negative binomial and binomial weights, respectively. For more details we refer to [4].

Statistical mechanics Multiplicative measures are used in this framework ([81]) under the names of macrocanonical and microcanonical ensembles (of particles). The multiplicative measures induced by Poisson, negative binomial and binomial distributions, with constant parameter functions, provide a mathematical setting for the three classical models of ideal gas, called Maxwell–Boltzmann, Bose–Einstein and Fermi–Dirac statistics.

CFP's on set partitions [9, 65] . This field stems from Pitman's study of combinatorial models of random set partitions, which developed from the Ewens sampling formula and Kingman's coalescence processes. We assume here that in the preceding setup for CFPs, particles are labeled by $1, \ldots, N$, so that the state space of a CFP is the set $\Omega_N = \left\{\pi_{[N]}\right\}$ of all partitions $\pi_{[N]}$ of the set $[N] = \{1, \ldots, N\}$. Denoting A_j the size of a cluster (block) $A_j \subseteq [N]$ each A_j has a weight $m_{|A_j|}$ which depicts the number of possible inner states of A_j, for example the number of shapes (in the plane or in space), colors, energy levels, and so forth. This says that to a set partition $\pi_{[n],k}$ with k given clusters $A_1, \ldots, A_k$ correspond $\prod_{j=1}^{k} m_{|A_j|}$ different states of the CFP considered. In a more general setting which encompasses a variety of models the weights m_j are allowed to be arbitrary non-negative numbers. By the known combinatorial relation, the distribution on $\Omega_{[N],k}$ given by the m_j's induces a multiplicative measure of cluster sizes $|A_j|$ on the set $\Omega_{N,k}$ of partitions of N into k (positive) summands.

3.1. Fluctuations near the limit shape of Young diagrams under a conservative measure

This section is based on a joint paper with Dirk Zeindler.

3.2. Randomization

We introduce in this section a probability measure $\mathbb{P}_t\left[\,\cdot\,\right]$ on $\dot{\cup}_{n\geq 1} \mathfrak{S}_n$, where $\dot{\cup}$ denotes the disjoint union, dependent on a parameter $t > 0$ with $\mathbb{P}_t\left[\,\cdot\,|\mathfrak{S}_n\right] = \mathbb{P}_n\left[\,\cdot\,\right]$ and consider the asymptotic behaviour of $w_n(x)$ with respect to $\mathbb{P}_t\left[\,\cdot\,\right]$ as $t \to 1$.

3.2.1. Grand canonical ensemble

Computations on $\mathfrak{S}_n$ can turn out to be difficult and many formulas can not be used to study the behaviour as $n \to \infty$. A possible solution to this problem is to adopt to a suitable randomization. This has been successfully introduced by [39] and used also by [13] as a tool to investigate combinatorial structures, and later applied in many contexts. The main idea of randomization is to define a one-parameter family of probability measures on $\dot{\cup}_{n\geq 1} \mathfrak{S}_n$ for which cycle counts turn out to be independent. Then one is able to study their behavior more easily, and ultimately the parameter is tuned in such a way that randomized functionals are distributed as in the non-randomized context. Let us see how to apply this in our work. We define

$$G_\Theta(t) = \exp\big(g_\Theta(t)\big) \tag{3.2.1}$$

with $g_\Theta(t)$ as in (3.0.5). If $G_\Theta(t)$ is finite for some $t > 0$, then for each $\sigma \in \mathfrak{S}_n$ let us define the probability measure

$$\mathbb{P}_t[\sigma] := \frac{1}{G_\Theta(t)} \frac{t^n}{n!} \prod_{k=1}^{n} \vartheta_k^{C_k}. \tag{3.2.2}$$

Lemma 3.5 shows that $\mathbb{P}_t$ is indeed a probability measure on $\dot\bigcup_{n\geq 1} \mathfrak{S}_n$. The induced distribution on cycle counts C_k can easily be determined.

Lemma 3.8. *Under $\mathbb{P}_t[\,\cdot\,]$ the C_k's are independent and Poisson distributed with*

$$\mathbb{E}_t[C_k] = \frac{\vartheta_k}{k} t^k.$$

Proof. From Pólya's enumeration theorem (Lemma 3.5) we obtain

$$\begin{aligned}
\mathbb{E}_t\left[e^{-sC_k}\right] &= \sum_{n\geq 0}\sum_{\sigma\in\mathfrak{S}_n} e^{-sC_k}\mathbb{P}_t[\sigma] = \frac{1}{G_\Theta(t)} \sum_{n\geq 0}\sum_{\sigma\in\mathfrak{S}_n} \frac{t^n}{n!} (\vartheta_k e^{-s})^{C_k} \prod_{\substack{j\leq n\\ j\neq k}} (\vartheta_j)^{C_j} \\
&= \frac{1}{G_\Theta(t)} \exp\left(\sum_{j=0}^{+\infty} \frac{\vartheta_j}{j} t^j\right) \exp\left(\left(e^{-s}-1\right)\frac{\vartheta_k}{k} t^k\right) \\
&= \exp\left(\left(e^{-s}-1\right)\frac{\vartheta_k}{k} t^k\right).
\end{aligned}$$

Analogously one proves the pairwise independence of cycle counts. □

Obviously the following conditioning relation holds:

$$\mathbb{P}_t[\,\cdot \mid \mathfrak{S}_n] = \mathbb{P}_n[\,\cdot\,].$$

A proof of this fact is easy and can be found for instance in [44, Equation (1)]. We note that $w_n(x)$ is $\mathbb{P}_t$-a.s. finite, since $\mathbb{E}_t[w_n(x)] < +\infty$. Now since the conditioning relation holds for all t with $G_\Theta(t) < +\infty$, one can try to look for t satisfying "$\mathbb{P}_n[\,\cdot\,] \approx \mathbb{P}_t[\,\cdot\,]$", which heuristically means that we choose a parameter for which permutations on $\mathfrak{S}_n$ weigh as most of the mass of the measure $\mathbb{P}_t$. We have on $\mathfrak{S}_n$

$$n = \sum_{j=1}^{\ell} \lambda_j = \sum_{k=1}^{n} kC_k.$$

A natural choice for t is thus the solution of

$$n = \mathbb{E}_t\left[\sum_{k=1}^{\infty} kC_k\right] = \sum_{k=1}^{\infty} \vartheta_k t^k. \tag{3.2.3}$$

which is guaranteed to exist if the series on the right-hand side is divergent at the radius of convergence (we will see this holds true for our particular choice of weights). We write $t = e^{-v_n}$ and use Lemma 3.6 in our case $\vartheta_k = \frac{k^\alpha + O(k^\beta)}{\Gamma(\alpha+1)}$ to obtain

$$n \stackrel{!}{=} (v_n)^{-\alpha-1} + O\big((v_n)^{-\beta-1}\big) \Longrightarrow v_n = (n^*)^{-1} + O\big((n^*)^{\beta-\alpha-1}\big) \tag{3.2.4}$$

with $n^* := n^{\frac{1}{1+\alpha}}$. We will fix this choice for the rest of the section.

3.2.2. Limit shape and mod-convergence

In order to derive our main results from the measure $\mathbb{P}_t$ we will use a tool developed by [55], the *mod-Poisson convergence*.

Definition 3.2.1. *A sequence of random variables $(Z_n)_{n\in\mathbb{N}}$ converges in the mod-Poisson sense with parameters $(\mu_n)_{n\in\mathbb{N}}$ if the following limit*

$$\lim_{n\to+\infty} \exp(\mu_n(1-e^{iu}))\mathbb{E}\left[e^{iuZ_n}\right] = \Phi(u)$$

exists for every $u \in \mathbb{R}$, and the convergence is locally uniform. The limiting function Φ is then continuous and $\Phi(0) = 1$.

This type of converge gives stronger results than a central limit theorem, indeed it implies a CLT (and other properties we will see below). Our goal will then be to prove the following

Proposition 3.2.2. *Let $x \geq 0$ be arbitrary and $x^* := xn^*$ with $n^* = n^{\frac{1}{1+\alpha}}$. Furthermore, let $t = e^{-v_n}$ with v_n as in* (3.2.4)*. Then the random variables $(w_n(x^*))_{n\in\mathbb{N}}$ converge in the mod-Poisson sense with parameters $\mu_n = (n^*)^\alpha w_\infty^{\mathbf{r}}(x) + o\big((n^*)^{\alpha/2}\big)$, where*

$$w_\infty^{\mathbf{r}}(x) := \frac{\Gamma(\alpha, x)}{\Gamma(\alpha+1)}. \tag{3.2.5}$$

$\Gamma(\alpha, x)$ is the upper incomplete Gamma function.

Proof. We have

$$\mathbb{E}_t\left[e^{isw_n(x^*)}\right] = \mathbb{E}_t\left[e^{is\sum_{\ell=\lfloor x^*\rfloor}^{\infty} C_\ell}\right] = \exp\left(\left(e^{is}-1\right)\sum_{j=\lfloor x^*\rfloor}^{\infty}\frac{\vartheta_j}{j}t^j\right). \tag{3.2.6}$$

This is the characteristic function of Poisson distribution. We thus obviously have mod-Poisson convergence with limiting function $\Phi(t) \equiv 1$. It remains to compute the parameter μ_n. Applying Lemma 3.6 for $x = 0$ and Lemma 3.7 for $x > 0$ together with (3.2.4) gives

$$\sum_{j=\lfloor x^* \rfloor}^{+\infty} \frac{j^{\alpha-1} + O\left(j^{\beta-1}\right)}{\Gamma(\alpha+1)} t^j = (n^*)^\alpha \frac{\Gamma(\alpha, x)}{\Gamma(\alpha+1)} + O\big((n^*)^\beta\big). \tag{3.2.7}$$

Since $\beta < \alpha/2$ by assumption, we deduce that $\lambda_n := (n^*)^\alpha w_\infty^{\mathbf{r}}(x) + o\big((n^*)^{\alpha/2}\big)$. This completes the proof. □

This yields a number of interesting consequences. In first place we can prove a CLT and detect the limit shape accordingly.

Corollary 3.2.3 (CLT and limit shape for randomization). *With the notation as above, we have as $n \to \infty$ with respect to $\mathbb{P}_t$*

$$\widetilde{w}_n^{\mathbf{r}}(x) := \frac{w_n(x^*) - (n^*)^\alpha w_\infty^{\mathbf{r}}(x)}{(n^*)^{\frac{\alpha}{2}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\sigma_\infty^{\mathbf{r}}(x))^2). \tag{3.2.8}$$

Furthermore the limit shape of $w_n(x)$ is given by $w_\infty^{\mathbf{r}}(x)$ (with scaling $A_n = (n^*)^\alpha$ and $B_n = n^*$, see (3.0.4)*). In particular, we can choose $\delta = 0$ in* (3.0.4).

Proof. The CLT follows immediately from [55, Prop. 2.4], but also can be deduced easily from (3.2.6) by replacing s by $s(n^*)^{-\alpha/2}$. It is also straightforward to show that $w_\infty^{\mathbf{r}}(x)$ is the limit shape. For a given $\epsilon > 0$, we choose $0 = x_0 < x_1 < \cdots < x_\ell$ such that $w_\infty^{\mathbf{r}}(x_{j+1}) - w_\infty^{\mathbf{r}}(x_j) < \epsilon/2$ for $1 \le j \le \ell - 1$ and $w_\infty^{\mathbf{r}}(x_\ell) < \epsilon/2$. It is now easy to see that for each $x \in \mathbb{R}^+$

$$|(n^*)^{-\alpha} w_n(x^*) - w_\infty^{\mathbf{r}}(x)| > \epsilon \Longrightarrow \exists j \text{ with } |(n^*)^{-\alpha} w_n(x_j^*) - w_\infty^{\mathbf{r}}(x_j)| > \epsilon/2.$$

Thus

$$\mathbb{P}_t \left[\sup_{x \ge 0} |(n^*)^\alpha w_n(x^*) - w_\infty^{\mathbf{r}}(x)| \ge \epsilon \right] \le \sum_{j=1}^{\ell} \mathbb{P}_t \left[|(n^*)^\alpha w_n(x_j^*) - w_\infty^{\mathbf{r}}(x_j)| \ge \epsilon/2 \right] \tag{3.2.9}$$

It now follows from (3.2.8) that each summand in (3.2.9) tends to 0 as $n \to \infty$. This completes the proof. □

Another by-product of mod-Poisson convergence of a sequence $(Z_n)_{n\in\mathbb{N}}$ is that one can approximate Z_n with a Poisson random variable with parameter μ_n, see [55, Prop. 2.5]. However in our situation this is trivial since $w_n(x^*)$ is already Poisson distributed.

As we are going to do in the next section, we are also interested in the behavior of increments and their joint behaviour.

Proposition 3.2.4. *For all $x, y \in \mathbb{R}$, $y > x$, set*

$$w_n(x,y) := w_n(x) - w_n(y) \ \text{ and } \ w^{\mathbf{r}}_\infty(x,y) := \frac{\Gamma(\alpha,x) - \Gamma(\alpha,y)}{\Gamma(\alpha+1)}.$$

Then

$$\widetilde{w}^{\mathbf{r}}_n(x,y) := \frac{w_n(x^*,y^*) - (n^*)^\alpha w^{\mathbf{r}}_\infty(x,y)}{(n^*)^{\frac{\alpha}{2}}\sqrt{w^{\mathbf{r}}_\infty(x,y)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \tag{3.2.10}$$

as $n \to \infty$ with $x^* = xn^{\frac{1}{\alpha+1}}$ and with $y^* = yn^{\frac{1}{\alpha+1}}$.

Furthermore, $\widetilde{w}^{\mathbf{r}}_n(x)$ and $\widetilde{w}^{\mathbf{r}}_n(x,y)$ are asymptotically independent.

Remark 3.9. *As we will see, the proof of independence relies on the independence of cycles coming from Lemma 3.8. Therefore it is easy to generalize the above result to more than two points.*

Proof. The proof of (3.2.10) almost the same as the proof of (3.2.8) and we thus omit it. Since

$$w_n(x,y) = \sum_{k=x^*}^{y^*-1} C_k \ \text{ and } \ w_n(y) = \sum_{k=y^*}^{\infty} C_k$$

and all C_k are independent, we have that $\widetilde{w}^{\mathbf{r}}_n(x)$ and $\widetilde{w}^{\mathbf{r}}_n(x,y)$ are independent for each $n \in \mathbb{N}$. Thus $\widetilde{w}^{\mathbf{r}}_n(x)$ and $\widetilde{w}^{\mathbf{r}}_n(x,y)$ are also independent in the limit. $\square$

3.2.3. Functional CLT

The topic of this section is to prove a functional CLT for the profile $w_n(x)$ of the Young diagram. Similar results were obtained in a different framework by [44, 24] on the number of cycle counts not exceeding $n^{\lfloor x \rfloor}$, and by [8] for Young diagrams confined in a rectangular box. We show

Theorem 3.2.5. *The process $\widetilde{w}^{\mathbf{r}}_n : \mathbb{R}^+ \to \mathbb{R}$ (see (3.2.8)) converges weakly with respect to $\mathbb{P}_t$ as $n \to \infty$ to a continuous process $\widetilde{w}^{\mathbf{r}}_\infty : \mathbb{R}^+ \to \mathbb{R}$ with $\widetilde{w}^{\mathbf{r}}_\infty(x) \sim \mathcal{N}(0, \sigma^{\mathbf{r}}_\infty(x))$ and independent increments.*

The technique we will exploit is quite standardized (see [44]). We remark that, unlike in this paper where the Ewens measure is considered, we do not obtain here a Brownian process, as the variance of $\widetilde{w}^{\mathbf{r}}_\infty(t) - \widetilde{w}^{\mathbf{r}}_\infty(s)$ for $r \geq s$ is more complicated than in the case of the Wiener measure. is represented We know from Proposition 3.2.4 the finite dimensional marginals of the process. More specifically we have for $x_\ell \geq x_{\ell-1} \geq \cdots \geq x_1 \geq 0$ that

$$(n^*)^{-\alpha/2}\big(w_n(x^*_\ell),\, w_n(x^*_{\ell-1}) - w_n(x^*_\ell),\, \ldots,\, w_n(x^*_1) - w_n(x^*_2)\big) \sim \mathcal{N}\left(\mathbf{0}, \Sigma'\right) \tag{3.2.11}$$

where Σ' is a diagonal matrix with

$$\Sigma'_{11} = w^{\mathbf{r}}_{\infty}(x_\ell) \text{ and } \Sigma'_{jj} = w^{\mathbf{r}}_{\infty}(x_{\ell-j+1}, x_{\ell-j+2}) \text{ for } j \geq 2.$$

Now all we need to show to complete the proof of Theorem 3.2.5 is the tightness of the process $\widetilde{w}^{\mathbf{r}}_n$. In order to do so, we will proceed similarly to [44], namely we will show that

Lemma 3.10. *We have for $0 \leq x_1 < x \leq x_2 < K$ with K arbitrary*

$$\mathbb{E}_t\left[(\widetilde{w}^{\mathbf{r}}_n(x) - \widetilde{w}^{\mathbf{r}}_n(x_1))^2(\widetilde{w}^{\mathbf{r}}_n(x_2) - \widetilde{w}^{\mathbf{r}}_n(x))^2\right] = O\left((x_2 - x_1)^2\right) \tag{3.2.12}$$

with $x^* := xn^{\frac{1}{\alpha+1}}$, $x_1^* := x_1 n^{\frac{1}{\alpha+1}}$ and $x_2^* := x_2 n^{\frac{1}{\alpha+1}}$.

Lemma 3.10 together with [11, Theorem 15.6] implies that the process $\widetilde{w}^{\mathbf{r}}_n$ is tight. This and the marginals in (3.2.11) prove Theorem 3.2.5.

Proof of Lemma 3.10. We define

$$E^* := \mathbb{E}_t\left[(\widetilde{w}^{\mathbf{r}}_n(x^*) - \widetilde{w}^{\mathbf{r}}_n(x_1^*))^2(\widetilde{w}^{\mathbf{r}}_n(x_2^*) - \widetilde{w}^{\mathbf{r}}_n(x^*))^2\right]. \tag{3.2.13}$$

Centering with $\mathbb{E}_t[w_n(\cdot)]$ and the independence of the cycle counts leads us to

$$\begin{aligned}
E^* &= \left(\sum_{k=x_1^*}^{x^*-1} (n^*)^{-\alpha}\frac{\theta_k}{k}t^k\right)\cdot\left(\sum_{k=x^*}^{x_2^*-1} (n^*)^{-\alpha}\frac{\theta_k}{k}t^k\right) \\
&\overset{Lem.\ 3.7}{\sim} \left(\frac{(n^*)^{-\alpha}}{\Gamma(\alpha+1)}\int_{x_1^*}^{x^*} t^{\alpha-1}e^{-t}\mathrm{d}t\right)\left(\frac{(n^*)^{-\alpha}}{\Gamma(\alpha+1)}\int_{x^*}^{x_2^*} t^{\alpha-1}e^{-t}\mathrm{d}t\right) \\
&= \left(\frac{\Gamma(\alpha,\, x_1) - \Gamma(\alpha,\, x)}{\Gamma(\alpha+1)}\right)\left(\frac{\Gamma(\alpha,\, x) - \Gamma(\alpha,\, x_2)}{\Gamma(\alpha+1)}\right) \\
&\sim O\left((x - x_1)(x_2 - x)\right) = O\left((x_2 - x_1)^2\right).
\end{aligned}$$

Here we have used the fact that $\Gamma(\alpha,\, \cdot)$ is a Lipschitz function and the assumption that $x_1 < x \leq x_2 < K$. □

3.3. Saddle point method

The aim of this section is to study the asymptotic behaviour of $w_n(x)$ with respect to $\mathbb{P}_n[\cdot]$ as $n \to \infty$ and to compare the results with the results in Section 3.2.

There are at least two approaches with which to tackle this problem: one is more probabilistic and was employed by [34] in their paper. The second one was first developed in [63] from the standard saddle point method.

The first method to study the asymptotic statistics of $w_n(x)$ with respect to $\mathbb{P}_n[\cdot]$ as $n \to \infty$ is the so called Khintchine method. We illustrate this method briefly with the normalisation constant h_n (see (3.0.2)). The first step is to write down a Khintchine's type representation for the desired quantity. For h_n this is given by

$$h_n = t^{-n} \exp\left(\sum_{k=1}^{n} \frac{\vartheta_k}{k} t^k\right) \mathbb{P}_t\left[\sum_{k=1}^{n} kC_k = n\right] \tag{3.3.1}$$

with $t > 0$ and $\mathbb{P}_t[\,\cdot\,]$ as in Section 3.2. The second step is to choose the free parameter t in such a way that $\mathbb{P}_t\left[\sum_{k=1}^{n} kC_k = n\right]$ gets large. Here one can choose t to be the solution of the equation $\sum_{k=1}^{n} \vartheta_k t^k = n$.

This argumentation is very close to the argumentation relying on complex analysis and generating functions. Indeed, it is easy to see that (3.3.1) is equivalent to

$$h_n = [t^n]\left[\exp\left(g_\Theta(t)\right)\right] \tag{3.3.2}$$

with $g_\Theta(t)$ as in (3.0.5). Furthermore, the choice of t is (almost) the solution of the saddle point equation $tg'_\Theta(t) = n$. We have of course to justify (3.3.2) (or (3.3.1)). But this follows immediately from the definition of h_n and Lemma 3.5.

We prefer at this point to work with the second approach. We begin by writing down the generating functions of the quantities we would like to study.

Lemma 3.11. *We have for $x \geq 0$ and $s \in \mathbb{R}$*

$$\mathbb{E}_n\left[\exp\left(-sw_n(x)\right)\right] = \frac{1}{h_n}[t^n]\left[\exp\left(g_\Theta(t) + (e^{-s} - 1)\sum_{k=\lfloor x \rfloor}^{\infty} \frac{\vartheta_k}{k} t^k\right)\right]. \tag{3.3.3}$$

Remark 3.12. *Although the expressions in Lemmas 3.11 and 3.13 hold in broader generality, starting from Subsection 3.3.1 we will calculate moment generating functions on the positive half-line, namely we can assume all parameters $s_1, \ldots, s_\ell$ etc to be non-negative, according to [19, Theorem 2.2].*

Proof. It follows from the definitions of $\mathbb{P}_n[\,\cdot\,]$ and $w_n(x)$ (see (3.0.3)) that

$$\begin{aligned} h_n\mathbb{E}_n\left[\exp\left(-sw_n(x)\right)\right] &= \frac{1}{n!}\sum_{\sigma \in \mathfrak{S}_n} \exp\left(-s\sum_{k=\lfloor x \rfloor}^{n} C_k\right)\prod_{k=1}^{n} \vartheta_k^{C_k} \\ &= \frac{1}{n!}\sum_{\sigma \in \mathfrak{S}_n} \prod_{k=1}^{\lfloor x \rfloor - 1} \vartheta_k^{C_k} \prod_{m=\lfloor x \rfloor}^{\infty} (\vartheta_k e^{-s})^{C_k} \end{aligned} \tag{3.3.4}$$

Applying now Lemma 3.5, we obtain

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} h_n \mathbb{E}_n \left[\exp\left(-s w_n(x)\right)\right] = \exp\left(\sum_{k=1}^{\lfloor x \rfloor - 1} \frac{\vartheta_k}{k} t^k + e^{-s} \sum_{k=\lfloor x \rfloor}^{\infty} \frac{\vartheta_k}{k} t^k\right) \tag{3.3.5}$$

$$= \exp\left(g_\Theta(t) + (e^{-s} - 1) \sum_{k=\lfloor x \rfloor}^{\infty} \frac{\vartheta_k}{k} t^k\right) \tag{3.3.6}$$

Equation (3.3.3) now follows by taking $[t^n]$ on both sides. □

We are also interested in the joint behaviour at different points of the limit shape. The results in Section 3.2 suggest that the increments of $w_n(x_{j+1}) - w_n(x_j)$ are independent for $x_\ell \geq x_{\ell-1} \geq \cdots \geq x_1 \geq 0$. It is thus natural to consider

$$\mathbf{w}_n(\mathbf{x}) = \left(w_n(x_\ell^*), w_n(x_{\ell-1}^*) - w_n(x_\ell^*), \ldots, w_n(x_1^*) - w_n(x_2^*)\right). \tag{3.3.7}$$

We obtain

Lemma 3.13. *We have for $\boldsymbol{x} = (x_1, \ldots, x_\ell) \in \mathbb{R}^\ell$ with $x_\ell \geq x_{\ell-1} \geq \cdots \geq x_1 \geq 0$ and $\boldsymbol{s} = (s_1, \ldots, s_\ell) \in \mathbb{R}^\ell$*

$$\mathbb{E}_n \left[\exp\left(-\langle \boldsymbol{s}, \boldsymbol{w}_n(\boldsymbol{x}) \rangle\right)\right] = \frac{1}{h_n} [t^n] \left[\exp\left(g_\Theta(t) + \sum_{j=1}^{\ell} (e^{-s_j} - 1) \sum_{k=\lfloor x_j \rfloor}^{\lfloor x_{j+1} - 1 \rfloor} \frac{\vartheta_k}{k} t^k\right)\right] \tag{3.3.8}$$

with the convention $x_{\ell+1} := +\infty$. The proof of this lemma is almost the same as for Lemma 3.11 and we thus omit it.

3.3.1. Log-n-admissibility

The approach with which we first addressed the study of the limit shape is derived from the saddle point method for approximating integrals in the complex plane. We want to introduce the definition of log-n-admissible function, generalizing the analogous concept introduced in [63]. We stress that here, in comparison to the definition of log- (or equivalently Hayman) admissibility used there, we consider a family of functions parametrized by n for which log-admissibility holds simultaneously. The definition is therefore a natural extension.

Definition 3.3.1. *Let $\left(g_n(t)\right)_{n \in \mathbb{N}}$ with $g_n(t) = \sum_{k=0}^{\infty} g_{k,n} t^k$ be given with radius of convergence $\rho > 0$ and $g_{k,n} \geq 0$. We say that $\left(g_n(t)\right)_{n \in \mathbb{N}}$ is* log-n-admissible *if there exist functions $a_n, b_n : [0, \rho) \to \mathbb{R}^+$, $R_n : [0, \rho) \times (-\pi/2, \pi/2) \to \mathbb{R}^+$ and a sequence $(\delta_n)_{n \in \mathbb{N}}$ s. t.*

Saddle-point *For each n there exists $r_n \in [0, \rho)$ with*

$$a_n(r_n) = n \tag{3.3.9}$$

Approximation *For all $|\varphi| \leq \delta_n$ we have the expansion*

$$g_n(r_n e^{i\varphi}) = g_n(r_n) + i\varphi a_n(r_n) - \frac{\varphi^2}{2} b_n(r_n) + R_n(r_n, \varphi) \tag{3.3.10}$$

where $R_n(r_n, \varphi) = o(\varphi^3 \delta_n^{-3})$.

Divergence $b_n(r_n) \to \infty$ *and* $\delta_n \to 0$ *as* $n \to \infty$.

Width of convergence *We have* $\delta_n^2 b_n(r_n) - \log b_n(r_n) \to +\infty$ *as* $n \to +\infty$.

Monotonicity *For all $|\varphi| > \delta_n$, we have*

$$\mathfrak{Re}\left(g_n(r_n e^{i\varphi})\right) \leq \mathfrak{Re}\left(g(r_n e^{\pm i\delta_n})\right). \tag{3.3.11}$$

The approximation condition allows us to compute the functions a and b exactly. We have

$$a_n(r) = r g_n'(r), \tag{3.3.12}$$

$$b_n(r) = r g_n'(r) + r^2 g_n''(r) \tag{3.3.13}$$

Clearly a_n and b_n are strictly increasing real analytic functions in $[0, \rho)$. The error in the approximation can similarly be bounded, so that

$$R_n(r, \varphi) = \varphi^3 O\left(r g_n'(r) + 3r^2 g_n''(r) + r^3 g_n'''(r)\right)$$

Having proved Lemma 3.11 we are now able to write down in a more explicit way generating functions. What we are left with is trying to extract the coefficients of the expansion given therein. This is the content of

Theorem 3.3.2. *Let $\left(g_n(t)\right)_{n\in\mathbb{N}}$ be* log*-n-admissible with associated functions a_n, b_n and constants r_n. Call*

$$G_n := [t^n] e^{g_n(t)}.$$

Then

(a) G_n has the asymptotic expansion

$$G_n = \frac{1}{\sqrt{2\pi}} (r_n)^{-n} b_n(r_n)^{-1/2} e^{g_n(r_n)} (1 + o(1)). \tag{3.3.14}$$

(b) Recall h_n defined in (3.0.2). *For the class of functions with weights as in* (3.0.8),

$$h_n = \frac{1}{\sqrt{2\pi(\alpha+1)n^{\frac{\alpha+2}{1+\alpha}}}} e^{2n^{\frac{\alpha}{1+\alpha}}}(1+o(1))$$

respectively as $n \to +\infty$.

Remark 3.14. *As it is explained in [36, Chapter VIII] it is possible to take into account more error terms in the expansion of g_n.*

Proof of Theorem 3.3.2. The proof is exactly the same as in [63, Prop. 2.2] and we thus give only a quick sketch of it, referring the reader to this paper for more details. As in the well-known saddle point method, we want to evaluate the integral

$$\frac{1}{2\pi i}\oint_\gamma \exp\left(g_n(z)\right)\frac{\mathrm{d}z}{z^{n+1}}.$$

We choose as contour the circle $\gamma := r_n e^{i\varphi}$ with $\varphi \in [-\pi, \pi]$. On $\varphi \in [-\delta_n, \delta_n]$ after changing to polar coordinates we can expand the function g as

$$\int_{-\delta_n}^{\delta_n} \exp\left(g_n(r) + i\varphi a_n(r) - \frac{\varphi^2}{2}b_n(r) + o(\varphi^3\delta_n^{-3}) - in\varphi\right)\mathrm{d}\varphi$$

We now choose r_n such that $a(r_n) = r_n g_n'(r_n) = n$ in order to cancel the linear terms in n. This allows us to approximate the integral on the minor arc with a Gaussian. One shows that away from the saddle point (so for $|\varphi| > \delta_n$) the contribution is exponentially smaller than on the minor arc and thus it can be neglected. □

We would like to emphasize also that it will be not always possible to solve the saddle point equation (3.3.9) exactly. However it is enough to find an r_n such that

$$a(r_n) - n = o\left(\sqrt{b(r_n)}\right) \tag{3.3.15}$$

holds.

3.3.2. Limit shape for polynomial weights

In this section we will derive the limit shape for Young diagrams for the class of measures given by the weights. We will not go into all the details to prove the log-n-admissibility for the most general case, but will try to give a precise overview of the main steps nonetheless. One important remark we have to make is that our parameter

s will not be fixed, but will be scaled and hence dependent on n. This comes from the fact that for a fixed s (3.3.9) becomes a fixed point equation whose solution cannot be given constructively, but has only an implicit form. We were not able to use this information for our purposes, and hence preferred to exploit a less general, but more explicit parameter to calculate asymptotics.

Limit shape

The main goal of this subsection is to prove that the weights (3.0.8) induce a sequence of log-n-admissible functions of which we can recover the asymptotics of $g_n(r_n)$. This will give us the limit shape of the Young diagram according to Theorem 3.3.2. More specifically

Theorem 3.3.3. *For the scaling $n^* = n^{\frac{1}{\alpha+1}}$, $x^* = xn^*$ and $s^* := s(n^*)^{-\alpha/2}$, define the functions*

$$w_\infty^{\mathbf{s}}(x) := \frac{\Gamma(\alpha, x)}{\Gamma(\alpha+1)} - (\alpha-1)\frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)},$$
$$\sigma_\infty^2(x) := \frac{\Gamma(\alpha, x)}{\Gamma(\alpha+1)} - \frac{\Gamma(\alpha+1, x)^2}{\Gamma(\alpha+1)\Gamma(\alpha+2)}$$

Then

$$\widetilde{w}_n^{\mathbf{s}}(x) := \frac{w_n(x^*) - (n^*)^\alpha w_\infty^{\mathbf{s}}(x)}{(n^*)^{\alpha/2}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma_\infty^2(x)\right).$$

In particular $w_\infty^{\mathbf{s}}(x)$ is the desired limit shape.

Remark 3.15. *Note that the limit shapes in the unrandomized and randomized cases are different, unlike [13].*

Theorem 3.3.4. *Define the cumulant generating function as*

$$\Lambda(s) := \log \mathbb{E}_n\left[e^{s\widetilde{w}_n(x)}\right] = \sum_{m\geq 1} q_m \frac{t^m}{m!}.$$

We then have for $m \geq 2$

$$q_m = \kappa_m(1 + o(1)). \tag{3.3.16}$$

with

$$\begin{aligned}\kappa_m \quad &= \quad [s^m]\left(\frac{(n^*)^\alpha}{\alpha}\left(1 - s^*\frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)}\right)^{-\alpha}\right. \\ &\quad + \left.\frac{\left(e^{-s^*}-1\right)}{\Gamma(\alpha+1)}\sum_{k\geq 0}\frac{\Gamma(\alpha+k,\, x)}{k!}\left(\frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)}s^*\right)^k\right).\end{aligned} \tag{3.3.17}$$

We can also determine the behavior of the increments of the function $w_n(\cdot)$. We will consider first the more general case and then give the example of the two-increment case (refer to (3.3.7) with $\ell = 2$).

Theorem 3.3.5. *(a) For $\ell \geq 2$ and $x_\ell \geq x_{\ell-1} \geq \cdots \geq x_1 \geq 0$, let*

$$\widetilde{\mathbf{w}}_n^{\mathbf{s}}(\mathbf{x}) = \left(\widetilde{w}_n^{\mathbf{s}}(x_\ell), \widetilde{w}_n^{\mathbf{s}}(x_{\ell-1}) - \widetilde{w}_n^{\mathbf{s}}(x_\ell), \ldots, \widetilde{w}_n^{\mathbf{s}}(x_1) - \widetilde{w}_n^{\mathbf{s}}(x_2)\right).$$

Set $x_{\ell+1} = +\infty$. For $1 \leq j < i < \ell$ we have that

$$\begin{aligned}\widetilde{w}_\infty^{\mathbf{s}}(x_i, x_j) &:= \lim_{n\to+\infty} \operatorname{Cov}\left(\widetilde{w}_n^{\mathbf{s}}(x_j) - \widetilde{w}_n^{\mathbf{s}}(x_{j+1}), \widetilde{w}_n^{\mathbf{s}}(x_i) - \widetilde{w}_n^{\mathbf{s}}(x_{i+1})\right) \\ &= \frac{\left(\Gamma(\alpha+1, x_i) - \Gamma(\alpha+1, x_{i+1})\right)\left(\Gamma(\alpha+1, x_j) - \Gamma(\alpha+1, x_{j+1})\right)}{\Gamma(\alpha+1)\Gamma(\alpha+2)}.\end{aligned}$$

Remark 3.16. *Let us comment briefly on Thm. 3.3.5. What we obtained in this result is most unexpected: cycle counts are asymptotically independent under very mild assumptions (see Lemma 3.4). The assumption of the lemma holds in our case as the growth of the parameters ϑ_n is algebraic. The fact that the increments depend on disjoint sets of cycles would have suggested the asymptotic independence of $w_n(y^*)$ from $w_n(x^*) - w_n(y^*)$. We are aware of the work of [6] handling this issue in the case of the Ewens sampling formula, in particular showing that partial sums of cycle counts need not converge to processes with independent increments. Our result extends this idea in the sense that it shows the explicit covariance matrix for a whole category of generating functions. It would be interesting to provide a heuristic explanation for this theorem.*

Log-n-admissibility

In order to determine the limit shape we would like to prove the log-n-admissibility of the function explicited in (3.3.3). To be more precise, what we have to prove is

Lemma 3.17. *Let $s \geq 0$, and recall $n^* = n^{\frac{1}{\alpha+1}}$, $s^* = s(n^*)^{\alpha/2}$. The function*

$$g_\Theta(t) + (e^{-s^*} - 1) \sum_{k=\lfloor x^* \rfloor}^{\infty} \frac{k^{\alpha-1} + O\left(k^{\beta-1}\right)}{\Gamma(\alpha+1)} t^k$$

is log-n-admissible for all $x \geq 0$, with $g_\Theta(t)$ as in (3.0.5) *and*

$$r_n := e^{-v_n} \tag{3.3.18}$$

with

$$v_n := (n^*)^{-1}\left(1 + s(n^*)^{-\alpha/2}\frac{\Gamma(\alpha+1, x)}{\Gamma(\alpha+2)}\right).$$

Proof of Lemma 3.17. **Saddle-point and approximation** We start first with the case $\beta = 0$. By doing so one obtains that

$$\begin{aligned}
a(r_n) &= \sum_{k=1}^{+\infty} \frac{k^\alpha}{\Gamma(\alpha+1)} e^{-kv_n} + (e^{-s^*}-1) \sum_{k=\lfloor x^* \rfloor}^{+\infty} \frac{k^\alpha}{\Gamma(\alpha+1)} e^{-kv_n} \\
&= (v_n)^{-\alpha-1} + O(1) + (e^{-s^*}-1)\frac{\Gamma(\alpha+1, x^* v_n)}{\Gamma(\alpha+1)} \\
&+ (e^{-s^*}-1) O\left(v_n^{-\alpha}\right) \qquad (3.3.19) \\
&= n\left(1 + (\alpha+1)s(n^*)^{-\alpha/2}\frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)} + O\left((n^*)^{-\alpha}\right)\right) \\
&+ n\left(-\frac{s}{(n^*)^{\alpha/2}} + O\left(\frac{s^2}{(n^*)^\alpha}\right)\right)\left(\frac{\Gamma(\alpha+1, x)}{\Gamma(\alpha+1)} + O\left((n^*)^{-\alpha/2}\right)\right) \\
&+ O\left(1 + v_n^{-\alpha}(n^*)^{-\alpha/2}\right) \\
&= n + O(n^*). \qquad (3.3.20)
\end{aligned}$$

We also have that

$$b(r_n) = O\left(\sum_{k=1}^{+\infty} \frac{k^{\alpha+1}}{\Gamma(\alpha+1)} e^{-kv_n}\right) \sim (\alpha+1)(n^*)^{\alpha+2} + O(n). \qquad (3.3.21)$$

Therefore (3.3.15) holds true for all α. In the case where β is turned on, we obtain by performing similar steps that

$$a(r_n) = n + O\left((n^*)^{\beta+1}\right).$$

Then (3.3.15) is satisfied if

$$\frac{\beta+1}{\alpha+1} < \frac{\alpha+2}{2(\alpha+1)} \iff \beta < \frac{\alpha}{2} \qquad (3.3.22)$$

which holds by assumption.

Divergence By the above calculations we set $\delta_n := (n^*)^{-\xi}$ with $\frac{\alpha+3}{3} < \xi < \frac{\alpha+2}{2}$. This position holds also in the case $\beta > 0$.

Monotonicity In the region $|\varphi| = o(1)$ we wish to show that

$$g\left(r_n e^{i\varphi}\right) = g(r_n)(1 + o(1)). \qquad (3.3.23)$$

First remember that $g_n\left(r_n e^{i \pm \delta_n}\right) = O\left((n^*)^\alpha\right)$ by Lemma 3.6. Thus here we have:

if $\varphi = o(v_n)$, then by a change of variable $t \rightsquigarrow (v_n - i\varphi)t$

$$\begin{aligned}&\sum_{k\geq \lfloor x^*\rfloor} \frac{k^{\alpha-1}}{\Gamma(\alpha+1)} e^{-k(v_n - i\varphi)} \\ &\sim \frac{(v_n - i\varphi)^{-\alpha}}{\Gamma(\alpha+1)} \int_x^{+\infty} t^{\alpha-1} e^{-t} \mathrm{d}t = \frac{\Gamma(\alpha, x)}{\Gamma(\alpha+1)} (v_n - i\varphi)^{-\alpha}\end{aligned}$$

which is asymptotic to $(n^*)^{\alpha}$. Considering the factor $e^{-s^*} - 1$ we obtain that the summand is negligible with respect to $\mathfrak{Re}\left(g(r_n e^{\pm i\delta_n})\right)$.

(b) If $\varphi \neq o(v_n)$, then

$$\begin{aligned}&\sum_{k\geq \lfloor x^*\rfloor} \frac{k^{\alpha-1}}{\Gamma(\alpha+1)} e^{-k(v_n - i\varphi)} \\ &\sim \frac{(v_n - i\varphi)^{-\alpha}}{\Gamma(\alpha+1)} \int_{x - ix\varphi n^* + o(1)}^{+\infty} t^{\alpha-1} e^{-t} \mathrm{d}t = \frac{\Gamma(\alpha, x - ix\varphi n^*)}{\Gamma(\alpha+1)} (v_n - i\varphi)^{-\alpha}\end{aligned}$$

and afterwards use the fact that $\Gamma(\alpha, x+iy) = O\left(y^{\alpha-1}\right)$ for $|y|$ large. Hence the RHS of (3.3.24) becomes

$$O\left((n^*)^{\alpha-1}\right)(v_n - i\varphi)^{1-\alpha} = O\left((n^*)^{\alpha-1}\varphi^{-1}\right)$$

As $\varphi \neq o(v_n)$, we obtain that $O\left((n^*)^{\alpha-1}\varphi^{-1}\right) = O\left((n^*)^{\alpha}\right) o(1)$ which is enough to show (3.3.23).

(c) To conclude we consider the case $|\varphi| > C$: the function $g_n\left(r_n e^{i\varphi}\right)$ is bounded there by a constant uniform in n, and then by bounding $g_n\left(r_n e^{i\varphi}\right)$ through its modulus we have

$$\mathfrak{Re}\left(g_n(r_n e^{i\varphi})\right) \leq \mathfrak{Re}\left(g(r_n e^{\pm i\delta_n})\right)\left(1 + O\left((n^*)^{-\alpha/2}\right)\right). \tag{3.3.24}$$

□

In order to show Thms. 3.3.3, 3.3.4 and 3.3.5 we need to prove first an auxiliary proposition.

Proposition 3.3.6. *For the scaling $n^* := n^{\frac{1}{\alpha+1}}$, $x^* := xn^*$ and $s^* := s(n^*)^{-\alpha/2}$*

$$\begin{aligned}\widetilde{w}_n^{\mathbf{s}}(x) &:= (e^{-s^*} - 1) \sum_{k\geq \lfloor x^*\rfloor} k^{\alpha-1} r_n^k \\ &= \left(-s(n^*)^{\alpha/2}\Gamma(\alpha, x) + \frac{s^2}{2}\Gamma(\alpha, x) - \frac{\Gamma(\alpha+1, x)^2}{\Gamma(\alpha+2)} s^2\right) + o(1)\end{aligned} \tag{3.3.25}$$

holds asymptotically as $n \to +\infty$.

Proof. We apply Lemma 3.7 with

$$\begin{aligned}
&f(t) := t^{\alpha-1}e^{-tv_n},\\
&z_n = xn^*,\ v_n = \frac{1}{n^*}\left(1 - s(n^*)^{-\alpha/2}\frac{\Gamma(\alpha+1,x)}{\Gamma(\alpha+2)}\right) \quad \text{and}\\
&z_n v_n = x - \frac{sx}{(n^*)^{\alpha/2}}\frac{\Gamma(\alpha+1,x)}{\Gamma(\alpha+2)}.
\end{aligned}$$

The first term of the expansion is

$$\begin{aligned}
&\left(e^{-s^*}-1\right)(n^*)^{\alpha}\,\Gamma(\alpha,x)\\
&= -s(n^*)^{\alpha/2}\Gamma(\alpha,x) + \frac{s^2}{2}\Gamma(\alpha,x) + o\,(1)
\end{aligned}$$

because $(s^*)^3(n^*)^{\alpha} = o\,(1)$. If $\beta > 0$ instead we obtain

$$\begin{aligned}
&\left(e^{-s^*}-1\right)(n^*)^{\alpha}\,\Gamma(\alpha,x)\\
&= -s(n^*)^{\alpha/2}\Gamma(\alpha,x) + \frac{s^2}{2}\Gamma(\alpha,x) + o\,(1)
\end{aligned}$$

To calculate the expansion up to a $O\,(1)$ term it is sufficient to consider for $k = 1$

$$\begin{aligned}
&\left(e^{-s^*}-1\right)(n^*)^{\alpha}\left(\frac{\Gamma(\alpha+1,x)}{\Gamma(\alpha+2)}sn^{-\beta}\right)\\
&= -\Gamma(\alpha+1,x)\frac{\Gamma(\alpha+1,x)}{\Gamma(\alpha+2)}s^2 + o\,(1)
\end{aligned}$$

This tells us that

$$\begin{aligned}
&\left(e^{-s^*}-1\right)\sum_{k\geq\lfloor x^*\rfloor} k^{\alpha-1}r_n^k\\
&= \left(-s(n^*)^{\alpha/2}\Gamma(\alpha,x) + \frac{s^2}{2}\Gamma(\alpha,x) - \frac{\Gamma(\alpha+1,x)^2}{\Gamma(\alpha+2)}s^2\right) + o\,(1) \qquad (3.3.26)
\end{aligned}$$

As for the remainder, we can find an a priori bound on the Bernoulli polynomials independent of n on $x \in [0,1]$. Furthermore,

$$\left(e^{-s^*}-1\right)f(\lfloor x^*\rfloor) = O\left(s(n^*)^{-\alpha/2}\right)(\lfloor xn^*\rfloor)^{\alpha-1}e^{-x+o(1)} = O\left(s(n^*)^{\frac{\alpha-2}{2}}\right),$$

which is small compared to the magnitude of the leading coefficient in s. Moreover

$$\begin{aligned}
&\int_{xn^*}^{+\infty} B_1(x'-\lfloor x'\rfloor)f'(x')\mathrm{d}x' \qquad (3.3.27)\\
&\leq C\int_{xn^*}^{+\infty}|f'(x')|\mathrm{d}x' = C\int_{xn^*}^{+\infty} e^{-x'v_n}(-v_n(x')^{\alpha-1}\\
&+(\alpha-1)(x')^{\alpha-2})\mathrm{d}x'. \qquad (3.3.28)
\end{aligned}$$

With the same substitution $x' := \frac{z_n}{a_0} y$ we can interchange limit and integral by the dominated convergence theorem to obtain

$$(3.3.28) = O\left(\left(\frac{z_n}{a_0}\right)^{\alpha/2} \Gamma(1+\alpha, a_0)\right).$$

Combining this with the first order expansion of $(e^{-s^*} - 1)$ we obtain

$$(3.3.27) = O\left(s(n^*)^{\alpha/2-\alpha/2}\right) = s\,O\left(1\right).$$

□

Proof of Thms. 3.3.3 and 3.3.4. To determine the behavior of G_n we would like to use Lemma 3.11. By (3.3.3)

$$\mathbb{E}_n\left[\exp\left(-s^* w_n(x^*)\right)\right] = \frac{1}{h_n}[t^n]\left[\exp\left(g_\Theta(t) + (e^{-s^*}-1)\sum_{k=\lfloor x^*\rfloor}^{+\infty} \frac{\vartheta_k}{k} t^k\right)\right].$$

We have shown that $g_n(t) = g_\Theta(t) + (e^{-s^*}-1)\sum_{k=\lfloor x^*\rfloor}^{+\infty} \frac{\vartheta_k}{k} t^k$ is log-n-admissible. Therefore Thm. 3.3.2 tells us how G_n behaves, and we have more precisely to recover three terms. In first place we collect the terms for the asymptotic of $e^{g_n(r_n)}$: one is

$$\begin{aligned}
&g_\Theta(r_n) \sim v_n^{-\alpha} = (n^*)^\alpha\left(1 - s(n^*)^{-\alpha/2}\frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)}\right)^{-\alpha} \\
&= (n^*)^\alpha + (n^*)^{\alpha/2}\frac{\alpha}{\Gamma(\alpha+2)}\left(s\Gamma(\alpha+1,\, x)\right) \\
&+ \frac{\alpha(\alpha+1)}{2\Gamma(\alpha+2)^2}\left(s\Gamma(\alpha+1,\, x)\right)^2 + O\left((n^*)^\beta\right)
\end{aligned}$$

given by Lemma 3.6. The other is

$$\sum_{k\geq\lfloor x^*\rfloor} \frac{k^{\alpha-1} + O\left(k^{\beta-1}\right)}{\Gamma(\alpha+1)} r_n^k = \frac{(n^*)^\alpha}{\Gamma(\alpha+1)} \sum_{k\geq 0} \frac{\Gamma(\alpha+k,\, x)}{k!}\left(s^* \frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)}(1+o\left(1\right))\right)^k \tag{3.3.29}$$

which we can approximate through Prop. 3.3.6. Secondly we obviously have

$$-n\log(r_n) = (n^*)^\alpha - (n^*)^{\alpha/2} s \frac{\Gamma(\alpha+1,\, x)}{\Gamma(\alpha+2)}.$$

Thirdly the behavior of $b(r_n)$ was determined in (3.3.21). All in all

$$e^{g(r_n)-n\log(r_n)} = \exp\left(2(n^*)^\alpha + s(n^*)^{\alpha/2}\left(\frac{\Gamma(\alpha,x)}{\Gamma(\alpha+1)} + (\alpha-1)\frac{\Gamma(\alpha+1,\,x)}{\Gamma(\alpha+2)}\right)\right.$$
$$+\frac{s^2}{2}\left(-\frac{1}{2\Gamma(\alpha+1)}\frac{\Gamma(\alpha+1,\,x)^2}{\Gamma(\alpha+2)} - \frac{\Gamma(\alpha+1,x)^2}{\Gamma(\alpha+2)} + \frac{\Gamma(\alpha,x)}{\Gamma(\alpha+1)}\right)$$
$$\left.+s^3 O\left((n^*)^{-3\alpha/2}\left(1+(n^*)^\alpha\right)\right)\right) \tag{3.3.30}$$

Theorem 3.3.2 yields the behavior of h_n, and the same theorem allows us to conclude plugging in (3.3.14) the expressions obtained in (3.3.21), (3.3.30) and h_n of (2) therein. It is also clear that $w_\infty^{\mathbf{s}}$ is the limit shape, in the same fashion the result followed in the proof of Corollary 3.2.3.

For cumulants what we have to do is considering the logarithm of the expansion $(r_n)^{-n}b(r_n)^{-1/2}e^{g_n(r_n)-n\log(r_n)}$. We claim that it suffices to consider simply the logarithm of the expression (3.3.30). In fact,

$$\log(b(r_n)) = \log\left(O\left((n^*)^{\alpha+2}\left(1-s^*\frac{\Gamma(\alpha,x)}{\Gamma(\alpha+1)}\right)^{-\alpha-2}\right)\right)$$
$$= C_1\log(n) + C_2\sum_{k\geq 0} s^k (n^*)^{\frac{k\alpha}{2}}$$

whilst each coefficient of s^k in $g_n(r_n) - n\log(r_n)$ is of order $(n^*)^{\frac{\alpha(2-k)}{2}}$ (compare (3.3.30)). This confirms that the main contribution stems from (3.3.30). □

Proof of Thm. 3.3.5. For multiple increments repeating the proof of Thm. 3.3.5 (b) tells us that for a vector $\mathbf{w}_n(\mathbf{x}^*)$ as in (3.3.7) with length $\ell > 2$ we can set

$$v_n := (n^*)^{-1}\left(1 + \frac{(n^*)^{-\alpha/2}}{\Gamma(\alpha+2)}\left(s_\ell\Gamma(\alpha+1,\,x_\ell)\right.\right.$$
$$\left.\left.+\sum_{k=1}^{\ell-1} s_{\ell-k}\left(\Gamma(\alpha+1,\,x_{\ell-1-k}) - \Gamma(\alpha+1,\,x_{\ell-k})\right)\right)\right).$$

We deduce from this that

$$g_\Theta(r_n) \sim v_n^{-\alpha} = (n^*)^\alpha - (n^*)^{\alpha/2}\frac{\alpha}{\Gamma(\alpha+2)}$$
$$\left(s_\ell\Gamma(\alpha+1,\,x_\ell) + \sum_{k=1}^{\ell-1}(s_{\ell-k-1}(\Gamma(\alpha+1,\,x_{\ell-k-1}) - \Gamma(\alpha+1,\,x_{\ell-k}))\right)$$
$$+\frac{\alpha(\alpha+1)}{2\Gamma(\alpha+2)^2}\left(s_\ell\Gamma(\alpha+1,\,x_\ell) + + \sum_{k=1}^{\ell-1}(s_{\ell-k-1}(\Gamma(\alpha+1,\,x_{\ell-k-1}) - \Gamma(\alpha+1,\,x_{\ell-k}))\right)^2$$
$$+o(1). \tag{3.3.31}$$

Since the coefficients of the form $\left(e^{-s_j^*}-1\right)\sum_{k=x_j^*}^{x_{j+1}^*-1}\frac{\vartheta_k}{k}r_n^k$ do not give a contribution to covariances, the mixed terms will stem from the expansion of the square in (3.3.31). In particular we see that the coefficient of $s_i s_j$, for $1 \le j < i < \ell$, is

$$\frac{\left(\Gamma(\alpha+1,\, x_i) - \Gamma(\alpha+1,\, x_{i+1})\right)\left(\Gamma(\alpha+1,\, x_j) - \Gamma(\alpha+1,\, x_{j+1})\right)}{2\Gamma(\alpha+1)\Gamma(\alpha+2)}.$$

□

3.3.3. Functional CLT for $w_n(\cdot)$

As in the randomized setting, a functional CLT can be obtained here too. Unlike the previous case though we do not have the independence of cycle counts, hence we will have to show the tightness of the fluctuations as in Sec. 3.2.3 in two steps (cf. [44]). The result we aim at is, precisely as before,

Theorem 3.3.7. *The process $\widetilde{w}_n^{\mathbf{s}} : \mathbb{R}^+ \to \mathbb{R}$ (see Thm. 3.3.3) converges weakly with respect to $\mathbb{P}_n$ as $n \to \infty$ to a continuous process $\widetilde{w}_\infty^{\mathbf{s}} : \mathbb{R}^+ \to \mathbb{R}$ with $\widetilde{w}_\infty^{\mathbf{s}}(x) \sim \mathcal{N}(0,\, (\sigma_\infty^{\mathbf{s}}(x))^2)$ and whose increments are* not *independent. The covariance structure is given in Thm. 3.3.5.*

Proof. We will proceed as in the proof of Thm. 3.2.5. Having shown already the behavior of the increments in Thm. 3.3.5 what we have to tackle now is their tightness. The proof's goal is again, analogoulsy as Lemma 3.10. However the evaluation of the LHS of (3.2.12) is more difficult this time; one possible approach is present in [24] and is based on Pólya's enumeration lemma and the calculation of factorial moments of cycle counts. We prefer rather to follow again [44]. We will proceed in two main steps.

i) We define for $0 < t < 1$ the measure $\mathbb{P}_t$ as in Section 3.2. By repeating the proof of [44, Lemma 2.1] we see that

$$\mathbb{P}_t\left[\sum kC_k = n\right] = t^n h_n e^{g_\Theta(t)}.$$

Mimicking Hansen's strategy one can also prove that for arbitrary functions $\Psi : \mathfrak{S} \to \mathbb{C}$, where $\mathfrak{S} := \cup_n \mathfrak{S}_n$ and $\Psi_n : \mathfrak{S}_n \to \mathbb{C}$ s. t. $\Psi_n = \Psi(C_1, \ldots, C_n, 0, 0, \ldots)$

$$\mathbb{E}_t\left[\Psi\right] e^{g_\Theta(t)} = \sum_{n\ge 1} t^n h_n \mathbb{E}\left[\Psi_n\right] + \Psi(0, 0, 0, \ldots). \tag{3.3.32}$$

ii) As a formal power series identity (3.3.32) holds for $|t| < 1$, thus we decide to set, for x_1, x_2 as in the assumptions,

$$\begin{aligned}
&\Psi(k_1, k_2, \ldots) := \\
&= \left(n^{-\gamma}\sum_{i=x_1^*+1}^{x^*}\left(k_i - \frac{\vartheta_i}{i}r_n^i\right)\right)^2\left(n^{-\gamma}\sum_{j=x^*+1}^{x_2^*}\left(k_j - \frac{\vartheta_j}{j}r_n^j\right)\right)^2 \\
&= \left(n^{-\gamma}\sum_{i=x_1^*+1}^{x^*} k_i - \frac{\vartheta_i}{i}t^i + n^{-\gamma}\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}\left(t^i - e^{-iv_n}\right)\right)^2 \\
&\left(n^{-\gamma}\sum_{j=x^*+1}^{x_2^*} k_j - \frac{\vartheta_j}{j}t^j + n^{-\gamma}\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}\left(t^j - e^{-jv_n}\right)\right)^2
\end{aligned}$$

for $\gamma > 0$ to be tuned appropriately later. We now calculate, using the independence of cycle counts under the randomized measure and the fact that $\mathrm{Var}_{\mathbb{P}_t}[C_i] = \mathbb{E}_t[C_i] = \frac{\vartheta_i}{i}t^i$,

$$\begin{aligned}
&\mathbb{E}_t[\Psi] = n^{-4\gamma}\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}t^i\right)\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}t^j\right) \\
&+n^{-4\gamma}\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}(t^i - e^{-iv_n})\right)^2\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}(t^j - e^{-jv_n})\right)^2 \\
&+2n^{-4\gamma}\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}t^i\right)\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}\left(t^i - e^{-iv_n}\right)\right) \\
&\cdot\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}(t^j - e^{-jv_n})\right)^2 \\
&+\ldots \\
&+2n^{-4\gamma}\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}t^i\right)\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}t^j\right)\left(\sum_{i=x_1^*+1}^{x^*}\frac{\vartheta_i}{i}\left(t^i - e^{-iv_n}\right)\right) \\
&\cdot\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}(t^j - e^{-jv_n})\right) \\
&=: G_\Theta^{(1)}(t,n) + G_\Theta^{(2)}(t,n) + \ldots + G_\Theta^{(9)}(t,n).
\end{aligned}$$

Let us define $g_a^b(z) := \sum_{j=a}^b \frac{\vartheta_j}{j}z^j$. From (3.3.32) we obtain

$$\mathbb{E}[\Psi_n] = \frac{1}{h_n}[t^n]\left(e^{g_\Theta(t)}G_\Theta^{(1)}(t,n)\right) + \ldots + \frac{1}{h_n}[t^n]\left(e^{g_\Theta(t)}G_\Theta^{(9)}(t,n)\right).$$

We therefore obtain several terms and we will analyze them one by one.

a) $\frac{1}{h_n}[t^n]\left(e^{g_\Theta(t)}G_\Theta^{(1)}(t,\,n)\right)$. One has

$$\begin{aligned}
&\frac{n^{-4\gamma}}{h_n}[t^n]\left(e^{g_\Theta(t)}g_{x_1^*+1}^{x^*}(t)g_{x^*+1}^{x_2^*}(t)\right)\\
&=\frac{n^{-4\gamma}}{h_n}[t^n]\left(e^{g_\Theta(t)+\log\left(g_{x_1^*+1}^{x^*}(t)\right)+\log\left(g_{x^*+1}^{x_2^*}(t)\right)}\right)
\end{aligned}\tag{3.3.33}$$

We want to apply the saddle-point method to the sequence of functions $g_n(t):=e^{g_\Theta(t)+\log\left(g_{x_1^*+1}^{x^*}(t)\right)+\log\left(g_{x^*+1}^{x_2^*}(t)\right)}$ to extract coefficients. Our first target is to show the log-n-admissibility. We consider again the radius $r_n := e^{-v_n}$ with $v_n := (n^*)^{-1}$. In this case as in (3.3.20)

$$\begin{aligned}
a(r_n)&=\sum_{k=1}^{+\infty}\frac{k^\alpha}{\Gamma(\alpha+1)}e^{-kv_n}+\frac{\sum_{x_1^*+1}^{x^*}\frac{k^\alpha}{\Gamma(\alpha+1)}e^{-kv_n}}{g_{x_1^*+1}^{x^*}(r_n)}+\frac{\sum_{x^*+1}^{x_2^*}\frac{k^\alpha}{\Gamma(\alpha+1)}e^{-kv_n}}{g_{x^*+1}^{x_2^*}(r_n)}=\\
&=(v_n)^{-\alpha-1}+O\left(1\right)+\frac{\sum_{x_1^*+1}^{x^*}\frac{k^\alpha}{\Gamma(\alpha+1)}e^{-kv_n}}{\sum_{x_1^*+1}^{x^*}\frac{k^\alpha-1}{\Gamma(\alpha+1)}e^{-kv_n}}+\frac{\sum_{x^*+1}^{x_2^*}\frac{k^\alpha}{\Gamma(\alpha+1)}e^{-kv_n}}{\sum_{x^*+1}^{x_2^*}\frac{k^\alpha-1}{\Gamma(\alpha+1)}e^{-kv_n}}\\
&=n+\frac{(v_n)^{-\alpha-2}C_{\alpha+1,\,x,\,x_1}}{v_n^{-\alpha-1}C_{\alpha,\,x,\,x_1}}+\frac{(v_n)^{-\alpha-2}C_{\alpha+1,\,x,\,x_2}}{v_n^{-\alpha-1}C_{\alpha,\,x,\,x_2}}\\
&=n+O\left(v_n^{-1}\right).
\end{aligned}$$

where $C_{\alpha+1,\,x,\,x_1}$, $C_{\alpha,\,x,\,x_1}$, $C_{\alpha+1,\,x,\,x_2}$ and $C_{\alpha,\,x,\,x_2}$ are constants independent of n. Very little changes also in the computions for $b(r_n)$ which lead to $b(r_n)=O\left((n^*)^{\alpha+2}\right)$, yielding the saddle point equation (3.3.15). As far as monotonicity is concerned, heuristically one can prove it using the fact that the order of $\log\left(g_{x_1^*+1}^{x^*}(t)\right)$ is smaller that that of the leading term $g_\Theta(t)$ (as one can already notice for example in the computations for $a(r_n)$ and $b(r_n)$ above). Since calculations are straightforward we omit them. Then by Thm. 3.3.2 one has that (recall that $h_n=[t^n]e^{g_\Theta(t)}$)

$$\begin{aligned}
&n^{-4\gamma}\left|\frac{1}{h_n}[t^n]e^{g_n(t)}\right|=\left|g_{x_1^*+1}^{x^*}(r_n)g_{x^*+1}^{x_2^*}(r_n)(1+o\left(1\right))\right|=\\
&\le Cn^{-4\gamma}\left|(v_n)^{-\alpha}\left(\Gamma(\alpha,\,x)-\Gamma(\alpha,\,x_1)\right)(v_n)^{-\alpha}\left(\Gamma(\alpha,\,x_2)-\Gamma(\alpha,\,x)\right)\right|\\
&\le Cn^{-4\gamma}(v_n)^{-2\alpha}\left|(x-x_1)\left(x_2-x\right)\right|\\
&=O\left((x-x_1)(x_2-x)\right)=O\left((x_2-x_1)^2\right)
\end{aligned}$$

provided that $n^{-4\gamma}(n^*)^{2\alpha} = O(1)$ iff $\gamma := \frac{\alpha}{2(\alpha+1)}$. We highlight that in this case n^γ is precisely the variance of the process (cf. Thm. 3.3.3). Here we have also used the fact that the incomplete Gamma function is continuous on a compact $[0, K]$ for some K large.

b) $\frac{1}{h_n}[t^n]\left(e^{g_\Theta(t)}G_\Theta^{(j)}(t,n)\right)$, $2 \le j \le 9$. We want to show that all these terms are $O\left((x_2-x_1)^2\right)$ as well. We take for example $G_\Theta^{(3)}(t,n) := \left(\sum_{j=x_1^*+1}^{x^*}\frac{\vartheta_j}{i}(t^j - e^{-jv_n})\right)^2\left(\sum_{j=x^*+1}^{x_2^*}\frac{\vartheta_j}{j}t^j\right)$. We define the auxiliary function $h_a^b(t) := \sum_{j=a}^{b}\frac{\vartheta_j}{j}(t^j - e^{-jv_n})$. We wish to apply again the saddle point method. In fact we decompose h as

$$h_{x_1^*+1}^{x^*}(t) = g_{x_1^*+1}^{x^*}(t) - \sum_{x_1^*+1}^{x^*}\frac{\vartheta_j}{j}e^{-jv_n}.$$

We now have

$$G_\Theta^{(3)}(t,n) = \left(g_{x_1^*+1}^{x^*}(t)\right)^2 g_{x^*+1}^{x_2^*}(t) - 2g_{x_1^*+1}^{x^*}(t)g_{x^*+1}^{x_2^*}(t)\left(\sum_{x_1^*+1}^{x^*}\frac{\vartheta_j}{j}e^{-jv_n}\right)$$
$$+\left(\sum_{x_1^*+1}^{x^*}\frac{\vartheta_j}{j}e^{-jv_n}\right)^2 g_{x^*+1}^{x_2^*}(t).$$

It is clear then that in the first-order asymptotics $G_\Theta^{(3)}$ (as well as all other terms involving $t^j - r_n^j$) will not give any contribution, because $G_\Theta^{(3)}(r_n, n) = 0$. We ask then ourselves if admissibility holds true for each one of these terms, but this is fairly easy because of the previous computations. Indeed we can start for example with the middle one. We have already shown in (a) that

$$n^{-4\gamma}\frac{1}{h_n}[t^n]\left(e^{g_\Theta(t)}g_{x_1^*+1}^{x^*}(t)g_{x^*+1}^{x_2^*}(t)\right)$$

is log-n-admissible and the term $\left(\sum_{x_1^*+1}^{x^*}\frac{\vartheta_j}{j}e^{-jv_n}\right)$ is a constant independent of t. Both the other two summands are log-n-admissible with $r_n = e^{-v_n} := e^{-1/n^*}$: calculations can be performed in the same fashion as (a) and since they are direct we skip them.

□

3.3.4. Large deviations estimates

We are able to prove large deviations estimates for $w_n(\cdot)$ thanks to our method as well. In fact, knowing the behavior of the Laplace transform enables us to compute the asymptotics of the Young diagram in the limit. More precisely, let σ_n be the limit variance as in Thm. 3.3.3. Define the normalized moment generating function and its logarithm as

$$\begin{aligned} M(s) &:= \mathbb{E}\left[\exp\left(s\frac{(w_n(x)-(n^*)^\alpha w_\infty^{\mathbf{s}}(x)}{\sigma_n}\right)\right], \\ \Lambda(s) &:= \log M(s). \end{aligned}$$

The strategy we adopt was first exploited in [63, Theorem 4.1], and relies on the fact that

Proposition 3.3.8. *There exist functions $\xi(n) = O\left((n^*)^\alpha\right)$, $\sigma(n) = O\left((n^*)^{\alpha/2}\right)$ such that for all $s = O\left(\sigma(n)\right)$ we obtain*

$$\Lambda(s) = \frac{s^2}{2} + O\left(\xi(n)\sigma(n)^{-3}\right)s^3.$$

It follows than that

$$\begin{aligned} \Lambda'(s) &= O\left(\xi(n)\sigma(n)^{-3}\right)s^2, \\ \Lambda''(s) &= O\left(\xi(n)\sigma(n)^{-3}\right)s. \end{aligned}$$

From this we derive

Proposition 3.3.9. *For all $a = O\left(\sigma_n\right)$ let $\delta := O\left(\xi(n)\sigma(n)^{-3}\right)$. Then we have*

$$\mathbb{P}\left[\left|\frac{(w_n(x)-(n^*)^\alpha w_\infty^{\mathbf{s}}(x)}{\sigma_n} - a\right| < \epsilon\right] = \left(1-\epsilon^{-2}(1+\delta)\right)\exp\left(-a^2/2 + O\left(\delta + \epsilon a\right)\right).$$

The error terms are absolute.

Proof. The proof can be performed analogously as [63], as we know that (3.3.30) holds. □

At this juncture we would like to apply our method to a simple but illustrative case.

3.3.5. An example: the case $g_\Theta(t) = (1-t)^{-1}$

We would like to begin by the easiest case, in other words to derive the limit shape for one point. We remark that here all our computations were performed using the function $g_\Theta(t) = t(1-t)^{-1}$. This does not affect the computations of the limit shape as it will "only" make a constant appear, which will be later simplified in all calculations.

Proposition 3.3.10. *For all $x \in \mathbb{R}^+$*

$$\frac{w_n(x\sqrt{n}) - \sqrt{n}e^{-x}}{n^{1/4}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, e^{-x}\left(1 - \frac{1}{2}e^{-x}(x+1)^2\right)\right)$$

In particular, the limit shape is $w_\infty^{\mathbf{s}}(x) := e^{-x}$ (cf. Thm. 3.3.3 plugging in $\alpha = 1$).

We now pass to the joint behavior of $(w_n(x_1), \dots, w_n(x_\ell))$ which can be recovered from

Proposition 3.3.11. *Let $\ell \in \mathbb{N}^+$. For all $x_1, \dots, x_\ell \in \mathbb{R}^+$, set $x_k^* := x_k\, n^{1/2}$; then we have*

$$\left(\frac{w_n\left(x_k^*\right) - n^{1/2}e^{-x_k}}{n^{1/4}}\right)_{k=1,\dots,\ell} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \Sigma\right)$$

with $\Sigma \in M_\ell(\mathbb{R})$ defined through

$$\begin{aligned} \Sigma_{k,k} = &\quad -\tfrac{1}{2}\, e^{-x_k}\left(e^{-x_k}{x_k}^2 + 2\, e^{-x_k}x_k + e^{-x_k} - 1\right) \\ \Sigma_{k,j} = &\quad -\tfrac{1}{2}\, e^{-x_k}\left(e^{-x_j}x_kx_j + e^{-x_j}x_k + e^{-x_j}x_j + e^{-x_j} - 1\right),\ j \neq k. \end{aligned} \tag{3.3.34}$$

Appendix A.

A.1. Gaussian bounds

Proof of (1.15) **and** (1.16)**.**
(1.15) For $t > a > 0$, $t + a > t - a$ and hence $t^2 - a^2 > (t-a)^2$,

$$\begin{aligned}\exp(a^2/2)P(|X| > a) &= 2\exp\left(a^2/2\right)P(X > a) = \\ &= 2\int_a^{+\infty} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2-a^2}{2}\right)\mathrm{d}t < \\ &< 2\int_a^{+\infty} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(t-a)^2}{2}\right)\mathrm{d}t = 1.\end{aligned}$$

Notice that the bound holds also at a=0.
(1.16) We have that the function

$$g(a) := 2P(X > a) - \frac{\exp\left(-a^2/2\right)}{\sqrt{2\pi}a}$$

is such that $g(1) > 0$, and its derivative

$$g'(a) = \frac{2}{\sqrt{2\pi}}\exp\left(-a^2/2\right)\left(\frac{1+a^2-a^3}{a^2}\right) < 0, \quad \forall\, a \geq 1.$$

Since $\lim_{a\to+\infty} g(a) = 0$, $g(a)$ is always non negative.

A.2. Bounds on Bessel functions

Here we will collect some of the bounds on the Bessel functions. These bounds are easy to derive but for completeness we provide a short proof for them.

Lemma A.1. *(a) For some constant $C > 0$ and $x > 0$*

$$\left|I_1^2(x) - I_0(x)I_2(x)\right| \geq Cx^2.$$

(b) Let $G(\cdot)$ be as in (2.9)*, then $G(x) \leq -C\log x$ for all $x \in [0,1]$, with $C > 0$ uniform in x.*

Proof. (a) Following [47] we have,

$$I_1^2(x) - I_0(x)I_2(x) = \frac{I_1^2(x)}{x}\left(x\frac{I_1'(x)}{I_1(x)}\right)' = \frac{I_1^2(x)}{x}\sum_{n\geq 1}\frac{4xj_{1,n}}{(x^2+j_{1,n}^2)^2}.$$

where we used the equality $\left(x\frac{I_1'(x)}{I_1(x)}\right)' = \sum_{n\geq 1}\frac{4xj_{1,n}}{(x^2+j_{1,n}^2)^2}$, $j_{i,n}$ being the n-th zero of $J_1(x)/x$ ([83]). Now using the identity $I_1(x) = (x/C)\prod_{n\geq 1}\left(1+\frac{x^2}{j_{1,n}^2}\right)$ [83, Page 498] we derive

$$\begin{aligned}
I_1^2(x) - I_0(x)I_2(x) &= \frac{I_1^2(x)}{x}\left(x\frac{I_1'(x)}{I_1(x)}\right)' \\
&= \frac{I_1^2(x)}{x}\frac{4xj_{1,1}}{(x^2+j_{1,1}^2)^2} + \frac{I_1^2(x)}{x}\sum_{n\geq 2}\frac{4xj_{1,n}}{(x^2+j_{1,n}^2)^2} \\
&> \frac{4I_1^2(x)j_{1,1}}{(x^2+j_{1,1}^2)^2} > C'x^2.
\end{aligned}$$

(b) By part (a) and the series expansion of Bessel functions ([1]) one can find a bound for $G(\cdot)$ as follows (γ is the Euler-Mascheroni constant):

$$\begin{aligned}
G(x) &\leq \frac{C}{x^2}(2I_1(x)K_1(x) + 2I_2(x)K_0(x) - 1) \\
&= \frac{C}{x^2}\left(2\left(\frac{x}{2}+\frac{x^3}{16}+O\left(x^4\right)\right)\left(\frac{1}{x}+\frac{x}{4}(-1+2\gamma-2\log 2+2\log x)\right.\right. \\
&+ O\left(x^3\log x\right)\Big) \\
&+ 2\left(\frac{x^2}{8}+O\left(x^4\right)\right)((-\gamma+\log 2-\log x) + \ + O\left(x^2\log x\right)\Big) - 1\Big) \\
&= \frac{C}{x^2}\left(1+\frac{x^2}{8}+O\left(x^3\right) + \frac{-1+2\gamma-2\log 2}{4}x^2 + \frac{-1+2\gamma-2\log 2}{32}x^4 + \right. \\
&+ O\left(x^2\log x\right) + \frac{x^2}{4}C + O\left(x^4\right) - \frac{x^2\log x}{4} - 1\Big) = -C\log x + C'.
\end{aligned}$$

Here C, C' denote positive constants that may vary from line to line. □

A.3. Euler Maclaurin formula with non integer boundaries

We prove in this section a slight extension of Euler Maclaurin formula, which allows to deal also with non-integer summation limits.

Theorem A.3.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function, $B_k(x)$ be the Bernoulli polynomials and $c < d$ with $c, d \in \mathbb{R}$. We then have for $p \in \mathbb{N}$*

$$\sum_{\lfloor c \rfloor \le k < d} f(k) = \int_c^d f(x)\,\mathrm{d}x - B_1(d - \lfloor d \rfloor) f(d) - B_1(c - \lfloor c \rfloor) f(c) \tag{A.3.1}$$
$$+ \sum_{k=1}^{p} (-1)^{k+1} \frac{B_{k+1}(d - \lfloor d \rfloor) f^{(k)}(d) - B_{k+1}(c - \lfloor c \rfloor) f^{(k)}(c)}{k!}$$
$$+ \frac{(-1)^{p+1}}{(p+1)!} \int_c^d B_{p+1}(x - \lfloor x \rfloor) f^{(p+1)}(x)\,\mathrm{d}x$$

Proof. The proof of this theorem follows the same lines as the proof of the Euler–Maclaurin summation formula with integer summation limits, see for instance [2, Theorem 3.1]. We give it here though for completeness. Our proof considers only the case $d \notin \mathbb{Z}$. The argumentation for $d \in \mathbb{Z}$ is completely similar. One possible definition of the Bernoulli polynomials is by induction:

$$B_0(y) \equiv 1, \tag{A.3.2}$$

$$B_k'(y) = kB_{k-1}(y) \ \text{ and } \int_0^1 B_k(y)\,\mathrm{d}y = 1 \text{ for } k \ge 1. \tag{A.3.3}$$

In particular, we have $B_1(y) = y - \frac{1}{2}$. We now have for $m \in \mathbb{Z}$

$$\begin{aligned}\int_m^{m+1} f(y)\,\mathrm{d}y &= \int_m^{m+1} B_0(y-m) f(y)\,\mathrm{d}y \\ &= [B_1(y-m)f(y)]|_{y=m}^{m+1} - \int_m^{m+1} B_1(y-m) f'(y)\,\mathrm{d}y \\ &= \frac{1}{2} f(m) + \frac{1}{2} f(m+1) - \int_m^{m+1} B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y.\end{aligned}$$

since $B_1(0) = -\frac{1}{2}$ and $B_1(1) = \frac{1}{2}$. We obtain

$$\sum_{k=\lfloor c \rfloor}^{\lfloor d \rfloor} f(k) = \int_{\lfloor c \rfloor}^{\lfloor d \rfloor} f(x)\,\mathrm{d}x + \frac{1}{2} f(\lfloor c \rfloor) + \frac{1}{2} f(\lfloor d \rfloor) + \int_{\lfloor c \rfloor}^{\lfloor d \rfloor} B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y.$$

Furthermore, we use

$$\int_{\lfloor d \rfloor}^{d} f(y)\,\mathrm{d}y = \frac{1}{2} f(\lfloor d \rfloor) + B_1(d - \lfloor d \rfloor) f(d) - \int_{\lfloor d \rfloor}^{d} B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y.$$

and get

$$\sum_{k=\lfloor c \rfloor}^{\lfloor d \rfloor} f(k) = \int_{\lfloor c \rfloor}^{d} f(x)\,\mathrm{d}x + \frac{1}{2} f(\lfloor c \rfloor) - B_1(d - \lfloor d \rfloor) f(d) + \int_{\lfloor c \rfloor}^{d} B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y.$$

The argumentation for replacing $\lfloor c \rfloor$ by c is similar. One gets

$$\sum_{\lfloor c \rfloor \leq k < d} f(k) = \int_c^d f(x)\,\mathrm{d}x - B_1(c - \lfloor c \rfloor) f(c) - B_1(d - \lfloor d \rfloor) f(d)$$
$$+ \int_c^d B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y.$$

The theorem now follows by successive partial integration of $\int_c^d B_1(y - \lfloor y \rfloor) f'(y)\,\mathrm{d}y$. $\square$

Notation

Notation	Meaning	Page
A^c	Complement of the set A	
$X \perp\!\!\!\perp Y$	X independent of Y	
$X \sim Y$	X and Y have the same distribution	
$\mathbb{R}^+$	$\mathbb{R} \setminus (-\infty, 0)$	
$\mathbb{N}^+$	$\mathbb{N} \setminus \{0\}$	
$\mathcal{N}(\mu, \sigma^2)$	Gaussian random variable of mean μ and variance σ^2	
$\mathbb{1}$	Indicator function of an event	
$\mathbb{I}$	Identity matrix	
$\delta_{x,y}$	Kronecker's delta	
$\langle \cdot, \cdot \rangle$	Scalar product	
$M_\ell(\mathbb{R})$	Set of $\ell \times \ell$ matrices with real coefficients	
$\equiv$	Constantly equal to	
$\mathcal{L}(X)$	Law of a random variable X	
$\mathcal{L}$	Convergence in distribution	
diam	Diameter of a set	
$\vec{e}$	Vector	
$a_n \sim b_n$	$\lim_{n \to +\infty} a_n / b_n = 1$	
$\dot{\cup}_n$	Disjoint union over n	
$\mathcal{B}(A)$	Borel sets of A	
$\partial\Lambda$	Boundary of Λ	14
Γ_Λ	Green's function of the SRW on Λ	14
Δ	Discrete Laplacian operator	15
τ_Λ	First exit time from Λ	14
$\overline{G}_N$	Modified Green's function	15
V_N	Discrete box of side-length N	16
G_N	Covariance matrix of the Membrane model	16
Δ_N^2	Discrete Bilaplacian operator with restriction	16
$\partial_2 V_N$	Double boundary of Λ	16

Notation	Meaning	Page
V_N^ℓ	Bulk of V_N	19
$\mathcal{H}_N(\eta)$	Set of high points of level η	20
φ_B	Conditional expectation of the center of a box	21
$\mathbf{Var}_B(\varphi_x)$	Conditional variance of φ_x	21
$I(h)$	Paley-Wiener integral	40
$I_k(\cdot)$	Modified Bessel function of order k	46
$T(a)$	Set of thick points of level a	46
$B(x, t)$		49
$\mathfrak{S}_n$	Set of permutations on n objects	65
$\vdash$	Cycle type	65
$w_n(\cdot)$	Young diagram	65
$C_k^{(n)}$, C_k	Cycle count of length k	66
$w_\infty(\cdot)$	Limit shape	67
$[\cdot]_n$	Extraction coefficient	69
$w_\infty^{\mathbf{r}}(x)$	Limit shape in the randomization	75
$\mathbf{w}_n(\mathbf{x})$	Increments vector	80
$w_\infty^{\mathbf{s}}(x)$	Limit shape in the saddle point method	83
$\sigma_\infty^2(x)$		83
$\widetilde{\mathbf{w}}_n(\mathbf{x})$	Rescaled increments	84

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables (Dover Books on Mathematics)*. Dover books on mathematics. Dover Publications, 1 edition, June 1965. ISBN 0486612724. URL http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0486612724. 98

[2] T.M. Apostol. *Introduction to analytic number theory*. Springer-Verlag, New York, 1984. 70, 99

[3] Tom M. Apostol. An elementary view of Euler's summation formula. *Amer. Math. Monthly*, 106(5):409–418, 1999. ISSN 0002-9890. doi: 10.2307/2589145. URL http://dx.doi.org/10.2307/2589145. 70

[4] R. Arratia, A.D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003. ISBN 3-03719-000-0. 68, 72

[5] M. Artin. *Algebra*. Pearson Education, 2011. ISBN 9780132413770. URL http://books.google.it/books?id=QsOfPwAACAAJ. 65

[6] Gutti Jogesh Babu, Eugenijus Manstavičius, and Vytas Zacharovas. Limiting processes with dependent increments for measures on symmetric group of permutations. In *Probability and number theory—Kanazawa 2005*, volume 49 of *Adv. Stud. Pure Math.*, pages 41–67. Math. Soc. Japan, Tokyo, 2007. 84

[7] M. Barlow and S. Taylor. Fractional dimension of sets in discrete spaces. *J.Phys. A: Math. Gen*, (22):2621–2626, 1989. 19

[8] Dan Beltoft, Cédric Boutillier, and Nathanaël Enriquez. Random Young diagrams in a Rectangular Box. *e-print*, December 2010. http://arxiv.org/abs/1008.0846. 77

[9] Nathanaël Berestycki and Jim Pitman. Gibbs Distributions for Random Partitions Generated by a Fragmentation Process. *Journal of Statistical Physics*, 127(2):381–418, 2007. ISSN 0022-4715. doi: 10.1007/s10955-006-9261-1. URL http://dx.doi.org/10.1007/s10955-006-9261-1. 72

[10] Volker Betz, Daniel Ueltschi, and Yvan Velenik. Random permutations with cycle weights. *Ann. Appl. Probab.*, 21(1):312–331, 2011. ISSN 1050-5164. doi: 10.1214/10-AAP697. URL `http://dx.doi.org/10.1214/10-AAP697`. 67

[11] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9. A Wiley-Interscience Publication. 78

[12] Marek Biskup and Oren Louidor. Extreme local extrema of two-dimensional discrete Gaussian free field. *ArXiv e-prints*, June 2013. 43

[13] Leonid V. Bogachev. Unified derivation of the limit shape for multiplicative ensembles of random integer partitions with equiweighted parts. *Random Structures & Algorithms*, pages n/a–n/a, 2014. ISSN 1098-2418. doi: 10.1002/rsa.20540. URL `http://dx.doi.org/10.1002/rsa.20540`. 68, 71, 73, 83

[14] E. Bolthausen, J. D. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *The Annals of Probability*, 29(4): 1670–1692, 2001. 28

[15] Maury Bramson, Jian Ding, and Ofer Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *ArXiv e-prints*, January 2013. URL `http://adsabs.harvard.edu/abs/2013arXiv1301.6669B`. 43

[16] Francesco Caravenna and Martin Borecki. Localization for $(1+1)$-dimensional pinning models with $(\nabla + \Delta)$-interaction. *Electron. Commun. Probab.*, 15:no. 48, 534–548, 2010. ISSN 1083-589X. doi: 10.1214/ECP.v15-1584. URL `http://ecp.ejpecp.org/article/view/1584`. 16

[17] Francesco Caravenna and Jean-Dominique Deuschel. Pinning and wetting transition for $(1+1)$-dimensional fields with Laplacian interaction. *Ann. Probab.*, 36(6): 2388–2433, 2008. ISSN 0091-1798. 16

[18] Francesco Caravenna and Jean-Dominique Deuschel. Scaling limits of $(1+1)$-dimensional pinning models with Laplacian interaction. *Ann. Probab.*, 37(3):903–945, 2009. doi: 10.1214/08-AOP424. 16

[19] Patrick Chareka. A finite-interval uniqueness theorem for bilateral Laplace transforms. *International Journal of Mathematics and Mathematical Sciences*, 2007:Article ID 60916, 6 p.–Article ID 60916, 6 p., 2007. URL `http://eudml.org/doc/54468`. 79

[20] Linan Chen and Dmitry Jakobson. Gaussian Free Fields and KPZ Relation in $\mathbb{R}^4$. *Annales Henri Poincaré*, 15(7):1245–1283, 2014. ISSN 1424-0637. doi: 10.1007/

s00023-013-0277-1. URL http://dx.doi.org/10.1007/s00023-013-0277-1. 43, 44, 45, 46, 50, 60

[21] A. Cipriani. High points for the membrane model in the critical dimension. *Electronic Journal of Probability*, 18, September 2013. 44, 47

[22] A. Cipriani and R. Subhra Hazra. Thick points for a Gaussian Free Field in 4 dimensions. *ArXiv e-prints*, July 2013. 43

[23] O. Daviaud. Extremes of the discrete two-dimensional Gaussian Free Field. *The Annals of Probability*, 34(3):962–986, 2006. 18, 19, 20, 44

[24] J. M. DeLaurentis and B. G. Pittel. Random permutations and Brownian motion. *Pacific J. Math.*, 119(2):287–301, 1985. ISSN 0030-8730. URL http://projecteuclid.org/getRecord?id=euclid.pjm/1102706156. 71, 77, 90

[25] Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni. Thick points for spatial Brownian motion: multifractal analysis of occupation measure. *Ann. Probab.*, 28 (1):1–35, 2000. doi: 10.1214/aop. 47

[26] Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni. Thick points for planar Brownian motion and the Erdős-Taylor conjecture on random walk. *Acta Math.*, 186(2):239–270, 2001. doi: 10.1007/BF02401841. 47

[27] Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni. Late points for random walks in two dimensions. *The Annals of Probability*, 34(1):219–263, 2006. 19

[28] Julien Dubédat. SLE and the free field: partition functions and couplings. *Journal of the American Mathematical Society*, 22(4):995–1054, 2009. 39, 41

[29] Bertrand Duplantier and Scott Sheffield. Liouville quantum gravity and KPZ. *Inventiones mathematicae*, 185(2):333–393, 2011. ISSN 0020-9910. doi: 10.1007/s00222-010-0308-1. URL http://dx.doi.org/10.1007/s00222-010-0308-1. 43, 44, 47, 48

[30] Richard Durrett, Boris L. Granovsky, and Shay Gueron. The Equilibrium Behavior of Reversible Coagulation-Fragmentation Processes. *Journal of Theoretical Probability*, 12(2):447–474, 1999. ISSN 0894-9840. doi: 10.1023/A:1021682212351. URL http://dx.doi.org/10.1023/A:1021682212351. 72

[31] Rick Durrett. *Probability: Theory and examples*. Cambridge University Press, 2011. 71

[32] Nicholas M. Ercolani and Daniel Ueltschi. Cycle structure of random permutations with cycle weights. *Random Structures & Algorithms*, 44(1):109–133, 2014.

ISSN 1098-2418. doi: 10.1002/rsa.20430. URL `http://dx.doi.org/10.1002/rsa.20430`. 66, 67, 68

[33] Paul Erdős and Joseph Lehner. The distribution of the number of summands in the partitions of a positive integer. *DUKE MATH . J*, 8(2):335–345, 1941. 71

[34] Michael M. Erlihson and Boris L. Granovsky. Limit shapes of Gibbs distributions on the set of integer partitions: the expansive case. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(5):915–945, 2008. ISSN 0246-0203. doi: 10.1214/07-AIHP129. URL `http://dx.doi.org/10.1214/07-AIHP129`. 69, 78

[35] W. J. Ewens. The sampling theory of selectively neutral alleles. *Theoret. Population Biology*, 3:87–112; erratum, ibid. 3 (1972), 240; erratum, ibid. 3 (1972), 376, 1972. ISSN 0040-5809. 68

[36] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, New York, NY, USA, 2009. ISBN 0521898064, 9780521898065. 70, 82

[37] G. Freiman, A. Vershik, and Y. Yakubovich. A Local Limit Theorem for Random Strict Partitions. *Theory of Probability & Its Applications*, 44(3):453–468, 2000. doi: 10.1137/S0040585X97977719. URL `http://epubs.siam.org/doi/abs/10.1137/S0040585X97977719`. 71

[38] Sacha Friedli and Yvan Velenik. Equilibrium Statistical Mechanics of Classical Lattice Systems: a Concrete Introduction. 2013. 16

[39] Bert Fristedt. The structure of random partitions of large integers. *Trans. Am. Math. Soc.*, 337(2):703–735, 1993. doi: 10.2307/2154239. 73

[40] Christophe Garban, June . 40, 43

[41] Giambattista Giacomin. Aspects of statistical mechanics of random surfaces. *Lectures at IHP Fall 2001*, 2001. 14

[42] V. L. Goncharov. Some facts from combinatorics. *Izvestia Akad. Nauk. SSSR*, 8: 3–48, 1944. 71

[43] Leonard Gross. Abstract Wiener spaces. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1*, pages 31–42. Univ. California Press, Berkeley, Calif., 1967. 40

[44] Jennie C. Hansen. A functional central limit theorem for the Ewens sampling formula. *J. Appl. Probab.*, 27(1):28–43, 1990. ISSN 0021-9002. 74, 77, 78, 90

[45] Fred M. Hoppe. The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.*, 25(2):123–159, 1987. ISSN 0303-6812. doi: 10.1007/BF00276386. URL `http://dx.doi.org/10.1007/BF00276386`. 68

[46] Xiaoyu Hu, Jason Miller, and Yuval Peres. Thick points of the Gaussian free field. *Ann. Probab.*, 38(2):896–926, 2010. doi: 10.1214/09-AOP498. 43, 44, 45, 48, 52, 53, 57

[47] C.M. Joshi and S.K. Bissu. Some inequalities of Bessel and modified Bessel functions. *J. Aust. Math. Soc., Ser. A*, 50(2):333–342, 1991. 98

[48] Jean-Pierre Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985. ISSN 0707-9109. 44

[49] J.P. Kahane. *Sur le chaos multiplicatif*. Prepublications mathematiques d'Orsay. Departement de mathematique, 1985. URL `http://books.google.ch/books?id=TUPEGwAACAAJ`. 42

[50] N.-G. Kang and N. Makarov. Gaussian free field and conformal field theory. *ArXiv e-prints*, January 2011. 39

[51] F.P. Kelly. *Reversibility and stochastic networks*. Wiley series in probability and mathematical statistics: Tracts on probability and statistics. J. Wiley, 1979. ISBN 9780471276012. URL `http://books.google.it/books?id=HhfvAAAAMAAJ`. 72

[52] Richard Kenyon. Dominos and the Gaussian free field. *Ann. Probab.*, 29(3):1128–1137, 2001. doi: 10.1214/aop. 39

[53] S. Kerov and A. Vershik. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables. *Dokl. Akad. Nauk SSSR*, 233:1024–1027, 1977. 71

[54] V. G. Knizhnik, A. M. Polyakov, and A. B. Zamolodchikov. Fractal structure of 2D-quantum gravity. *Modern Phys. Lett. A*, 3(8):819–826, 1988. ISSN 0217-7323. doi: 10.1142/S0217732388000982. URL `http://dx.doi.org/10.1142/S0217732388000982`. 43

[55] Emmanuel Kowalski and Ashkan Nikeghbali. Mod-Poisson convergence in probability and number theory. *Int. Math. Res. Not. IMRN*, 2010(18):3549–3587, 2010. ISSN 1073-7928. 75, 76

[56] N. Kurt. *Entropic repulsion for a Gaussian membrane model in the critical and supercritical dimension*. PhD thesis, University of Zurich, 2008. 17, 46

[57] N. Kurt. Maximum and entropic repulsion for a Gaussian membrane model in the critical dimension. *The Annals of Probability*, 37(2):687–725, 2009. 16, 19, 22, 23, 27, 28, 44

[58] Noemi Kurt. Entropic repulsion for a class of Gaussian interface models in high dimensions. *Stochastic Processes Appl.*, 117(1):23–34, 2007. doi: 10.1016/j.spa.2006.05.011. 16, 35, 44

[59] G. Lawler and V. Limic. *Random walk: a modern introduction*. Cambridge University Press, Cambridge, 2010. 15, 16

[60] B.F Logan and L.A Shepp. A variational problem for random Young tableaux. *Advances in Mathematics*, 26(2):206–222, 1977. ISSN 0001-8708. doi: 10.1016/0001-8708(77)90030-5. URL `http://www.sciencedirect.com/science/article/pii/0001870877900305`. 71

[61] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. ISBN 0-19-853489-2. With contributions by A. Zelevinsky, Oxford Science Publications. 69

[62] T. Madaule. Maximum of a log-correlated Gaussian field. *ArXiv e-prints*, July 2013. URL `http://adsabs.harvard.edu/abs/2013arXiv1307.1365M`. 43

[63] Kenneth Maples, Ashkan Nikeghbali, and Dirk Zeindler. The number of cycles in a random permutation. *Electron. Commun. Probab.*, 17:no. 20, 1–13, 2012. ISSN 1083-589X. doi: 10.1214/ECP.v17-1934. URL `http://ecp.ejpecp.org/article/view/1934`. 68, 78, 80, 82, 94

[64] P. Mörters, Y. Peres, O. Schramm, and W. Werner. *Brownian Motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010. ISBN 9780521760188. URL `http://books.google.ch/books?id=e-TbA-dSrzYC`. 55

[65] Jim Pitman. *Combinatorial stochastic processes*, volume 1875. Springer-Verlag, 2006. 72

[66] Boris Pittel. On a Likely Shape of the Random Ferrers Diagram. *Advances in Applied Mathematics*, 18(4):432–488, 1997. ISSN 0196-8858. doi: 10.1006/aama.1996.0523. URL `http://www.sciencedirect.com/science/article/pii/S0196885896905230`. 71

[67] Boris Pittel. On the distribution of the number of Young tableaux for a uniformly random diagram. *Advances in Applied Mathematics*, 29(2):184–214, 2002. ISSN 0196-

8858. doi: 10.1016/S0196-8858(02)00009-X. URL http://www.sciencedirect.com/science/article/pii/S019688580200009X. 71

[68] G. Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen. *Acta Mathematica*, 68:145–254, 1937. 69

[69] Rémi Rhodes and Vincent Vargas. Gaussian multiplicative chaos and applications: a review. *ArXiv e-prints*, May 2013. URL http://arxiv.org/pdf/1305.6221v1.pdf. 42, 43, 44, 47

[70] B. Rider and B. Virág. The noise in the circular law and the Gaussian free field. *ArXiv Mathematics e-prints*, June 2006. 39

[71] H. Sakagawa. Entropic repulsion for a Gaussian lattice field with certain finite range interactions. *J. Math. Phys.*, 44(7):2939–2951, 2003. 16

[72] Oded Schramm and Scott Sheffield. A contour line of the continuum Gaussian free field. *Probability Theory and Related Fields*, pages 1–34, 2012. ISSN 0178-8051. doi: 10.1007/s00440-012-0449-9. URL http://dx.doi.org/10.1007/s00440-012-0449-9. 39

[73] Scott Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Relat. Fields*, 139(3-4):521–541, 2007. doi: 10.1007/s00440-006-0050-1. 40, 44

[74] Daniel Stroock. Abstract Wiener space, revisited. *Comm. Stoch. Anal.*, 2(1):145–151, 2008. 61

[75] Daniel Stroock. *Probability Theory: An Analytic View*. Cambridge University Press, 2010. ISBN 9780521761581. URL http://books.google.ch/books?id=IvSSLeXpq3sC. 40, 42, 45, 50

[76] M. Szalay and P. Turán. On some problems of the statistical theory of partitions with application to characters of the symmetric group. III. *Acta Mathematica Academiae Scientiarum Hungarica*, 32(1-2):129–155, 1978. ISSN 0001-5954. doi: 10.1007/BF01902208. URL http://dx.doi.org/10.1007/BF01902208. 71

[77] Alain-Sol Sznitman. *Topics in occupation times and Gaussian free fields*. European Mathematical Society, 2012. 23

[78] H. N. V. Temperley. Statistical mechanics and the partition of numbers. II. The form of crystal surfaces. *Proc. Camb. Phil. Soc.*, 48:683–697, 1952. 71

[79] Y. Velenik. Localization and delocalization of random interfaces. *ArXiv Mathematics e-prints*, September 2005. 14

[80] A. Vershik and Yu. Yakubovich. The limit shape and fluctuations of random partitions of naturals with fixed number of summands. *Mosc. Math. J.*, 1(3):457–468, 472, 2001. ISSN 1609-3321. 71

[81] A. M. Vershik. Statistical mechanics of combinatorial partitions, and their limit configurations. *Funktsional. Anal. i Prilozhen.*, 30(2):19–39, 96, 1996. ISSN 0374-1990. doi: 10.1007/BF02509449. URL http://dx.doi.org/10.1007/BF02509449. 71, 72

[82] A.M. Vershik and Yu.V. Yakubovich. Asymptotics of the Uniform Measures on Simplices and Random Compositions and Partitions. *Functional Analysis and Its Applications*, 37(4):273–280, 2003. ISSN 0016-2663. doi: 10.1023/B:FAIA.0000015578.02338.0e. URL http://dx.doi.org/10.1023/B:FAIA.0000015578.02338.0e. 71

[83] G. N. Watson. *A treatise on the theory of Bessel functions*. Cambridge University Press, 1944. 98

[84] Yu. Yakubovich. Asymptotics of random partitions of a set. *Journal of Mathematical Sciences*, 87(6):4124–4137, 1997. ISSN 1072-3374. doi: 10.1007/BF02355807. URL http://dx.doi.org/10.1007/BF02355807. 71